

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:56 ; Search time 7228 Seconds
(without alignments)
4113.631 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPITAYAAQTRGLGLCIIT.....PAIIPDREVLVREFDEMEEC 686

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US09930591/runat_15092004_164646_7096/app.query.fasta_1.839
-DB=GenEmbl -CFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pco -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09930591 @CGN_1_1_4545 @runat_15092004_164646_7096 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARM TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sv.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3618	100.0	2061	6	AX441176	Sequence AX441176
2	3618	100.0	2061	6	AX467113	Sequence AX467113
3	3574	98.8	5360	6	I06434	Sequence I06434
4	3574	98.8	5360	6	I09328	Sequence I09328
5	3574	98.8	6785	6	I06440	Sequence I06440
6	3574	98.8	6785	6	I09329	Sequence I09329
7	3574	98.8	7310	6	AR118696	Sequence AR118696
8	3574	98.8	7310	6	I09331	Sequence I09331
9	3574	98.8	7310	14	HPCPOLXP	Sequence HPCPOLXP
10	3574	98.8	9185	6	I08294	Sequence I08294
11	3574	98.8	9185	6	BD091382	Sequence BD091382
12	3574	98.8	9379	6	AR166930	Sequence AR166930
13	3574	98.8	9379	6	AR301300	Sequence AR301300
14	3574	98.8	9401	6	AR176483	Sequence AR176483
15	3574	98.8	9401	6	E66593	Sequence E66593
16	3574	98.8	9401	6	I71894	Sequence I71894
17	3574	98.8	9401	6	I81885	Sequence I81885
18	3574	98.8	9401	6	BD080334	Sequence BD080334
19	3574	98.8	9401	14	HPCPLYPRE	Sequence HPCPLYPRE
20	3574	98.8	9609	12	AF387805	Sequence AF387805
21	3574	98.8	9646	12	AF387806	Sequence AF387806
22	3574	98.8	9693	12	AF387807	Sequence AF387807
23	3571	98.7	6785	6	AR118692	Sequence AR118692
24	3570	98.7	8316	6	AR118703	Sequence AR118703
25	3570	98.7	9185	6	AR118722	Sequence AR118722
26	3570	98.7	9185	6	AR118723	Sequence AR118723
27	3570	98.7	9618	14	AF271632	Sequence AF271632
28	3568	98.6	6299	6	AX164584	Sequence AX164584
29	3565	98.5	2058	6	AR404933	Sequence AR404933
30	3565	98.5	2058	6	AR408362	Sequence AR408362
31	3565	98.5	2058	6	AX395309	Sequence AX395309
32	3565	98.5	2058	6	AX454818	Sequence AX454818
33	3565	98.5	8987	6	AR118728	Sequence AR118728
34	3565	98.5	9609	12	AF387808	Sequence AF387808
35	3555	98.3	5360	6	AR118686	Sequence AR118686
36	3551	98.1	9646	6	BD069982	Sequence BD069982
37	3551	98.1	9646	6	BD069982	Sequence BD069982
38	3551	98.1	12980	6	AF009606	Sequence AF009606
39	3551	98.1	12980	6	AR110831	Sequence AR110831
40	3551	98.1	12980	6	BD069985	Sequence BD069985
41	3550	98.1	9395	14	AF511950	Sequence AF511950
42	3550	98.1	9502	6	E08263	Sequence E08263
43	3550	98.1	9502	6	E08264	Sequence E08264
44	3550	98.1	9502	14	HPCHCJ1	Sequence HPCHCJ1
45	3547	98.0	9610	14	HEC278830	Sequence HEC278830

ALIGNMENTS

RESULT 1

AX441176 AX441176 2061 bp DNA linear PAT 28-JUN-2002

LOCUS Sequence 16 from Patent WO0213855.

DEFINITION AX441176

ACCESSION AX441176

VERSION AX441176.1 GI:21665758

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM artificial sequences.

REFERENCE 1

AUTHORS Sallberg, M. and Huitgren, C.

TITLE Vaccines containing ribavirin and methods of use thereof

JOURNAL Patent: WO 0213855-A 16 21-FEB-2002;

TRIPEP AB (SE)

FEATURES

source Location/Qualifiers

1..2061

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="Hepatitis C virus NS3/4A coding region"

ORIGIN

Alignment Scores:

Pred. No.: 1.41e-203 Length: 2061

Score: 3618.00 Matches: 686

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AX441176 (1-2061)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20

Db 1 ATGGCGCTATCAGGGCTATGCCACACAGAGGGGCGCTTTGGGATGCATAATCACC 60

Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40

Db 61 AGCTTGACCGCGCGGCAAAACACAGGTGGGGTGGAGTTCCAGATCGTGTCACCTGCT 120

Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60

Db 121 GCCCAGACTTCTGGCAACTGCATTAAACGGGGTGTGTGGAGTGTCTACCATGGAGCC 180

Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnVetTyrThrAsnValAsp 80

Db 181 GGAAACAGGACCAATGGCTCACCTAAGGTCCTGTATCCAGATGTACACCAATGTGGAC 240

Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlnValaArgSerLeuThrProCysThrCys 100

Db 241 CAAGACCTGTAGCTGGCGCTGCCCAAGGTGGCGCTCATTAACCATGCACTTGC 300

Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

Db 301 GCCTCTCGGACCTTACCTGGTCACGAGCAGCGCGATGTCATTCTGTGCGCGGACG 360

Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140

Db 361 GGTGATGGCAGGGCAGCGCTGCTTCGCGCGCGCTATCTCTTACTTGAAGGCTCCTCG 420

Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160

Db 421 GGAGGCGCTCTGCTGTGCGCGCAGGACATGCGTAGGCATATTTCAGAGCGCGGTATGC 480

Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGluThrThrMet 180

Db 481 ACCCGTGGAGTGGTAAAGCGGTGGACTTCATCCCGTAGAGGCTTAGAGCAACCATG 540

Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200

Db 541 AGTCCCGGCTGTCTCAGACCACTCTCCACCACAGCAGTGGCCCCAGAGCTACCAAGTG 600

Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220

601 GCCCACTGCATGCTCCACCGGCGAGCGTAAGACCAACCAAGGTCCCGCGCATACGCA 660

Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

Db 661 GCTCAGGCGCTACAAGGTGCTGCTCAACCCCTCCGTTCGCAACAATGGGCTTTGGT 720

Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260

Db 721 GCTTACATGTCACAGGCCCATGGATTGATCCTTAACATCAGACTGGGGTCAGACAAAT 780

Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280

Db 781 ACTACTGGCAGCGGATCAGCTATTCCACCTACGCAAGTTCTTGGCCGACGGGGGTGT 840

Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300

Db 841 TCAGGGGTGCTTATGACATATATATTTGTGACAGTGCCACTCCACGGATGCAACATCC 900

Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320

Db 901 ATCTTGGGCATTGGCACTGCTCTTGACCAAGCAGACCGCGGGCGAGACTGACTGTG 960

Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal 340

Db 961 CTCGCCACCGCTACCCCTCCGGGCTCCGTCACCTGTCCTCCATCTTAACATCGAGAGGT 1020

Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360

Db 1021 GCTCTGTCCACTACGGGAGAGATCCCTTTTATGCAAGGCTATTCCCTTGAAGCAAT 1080

Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380

Db 1081 AAGGGGGGAGACATCTCATCTTCTGCACCTCAAGAAGAAGTCCGACGAGCTCGCGCA 1140

Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400

Db 1141 AACTGTGCGGTGGGGCTCAATCGCTGCTTACTACCGGGCTTGTATGTCCGTC 1200

Qy 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420

Db 1201 ATCCCGACAGTGTGCTGCTGTCGCAACTGACGCCCTCATGACCGGCTTTTACC 1260

Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

Db 1261 GCGCACTTCGATTTCGTTAGACTGCAACAGTGTGTACCCAGACAGCTGCATTCAGC 1320

Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460

Db 1321 CTTGACCTTACCTTCCACCATGAGACAATCAGCTTCCCGAGATGCTGTCTCCGCTACT 1380

Qy 461 GlnArgArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly 480

Db 1381 CAACGTCGGGTAGACTTGGCAGAGGAAGCCAGCATCTACAGATTGTGGCACCGGG 1440

Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

Db 1441 GAGCGTCTCTTGGCATGTTTACTGCTGTCTCTCTGCGAGTGTATGACGCGGTGT 1500

Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520

Db 1501 GCTTGGTATGAGCTTACGCGCGCAGACCATCTTGAATTTGGGAGGGCGCTTACGGGCTC 1560

Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540

Db 1561 CCGGACTTCCGCTGTGCCAAGACCATCTTGAATTTGGGAGGGCGCTTACGGGCTC 1620

Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560

Db 1621 ACCCAGATAGACGGCCACTTCTTATCCAGACAAGCAGAGTGGGGAACCTTCCCTAT 1680

Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580

```

Db      1681 CTGGTAGCGTACCAAGCCACCGCTGTGCGGTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740
QY      581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      1741 CAGATGTGGAGTGTGTATCGGTCTCAAGCCACCTTCATGGGCCACACCTCTGTGTA 1800
QY      601 TyrArgLeuGlyAlaValGlnAenGluValThrLeuThrHisProValThrLysTyrIle 620
Db      1801 TATAGACTGGCGCTGCTCCAGAAATGAAGTCAAGCTGACGCCACCCAGTCACCAAGTATATC 1860
QY      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db      1861 ATGACATGTATGTGCGCTGACCTGGAGGTGCTGACAGTAGTACTGGTGTCTGTTGGCGGC 1920
QY      641 ValLeuAlaLeuAlaValCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      1921 GTTCTGGCTGCTTTGGCCGCGTATTGCCGTATCCACAGGTGCGTGTGTCATAGTAGTAGG 1980
QY      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuThrArgGluPhe 680
Db      1981 ATTGTCTTGTCCGGAAGCGCGCAATCATCCGACAGGGAAGTCTCTTACCGGAGTTC 2040
QY      681 AspGluMetGluGluCys 686
Db      2041 GATGAATGGAGAGTGC 2058

RESULT 2
AX467113 2061 bp DNA linear PAT 17-JUL-2002
LOCUS AX467113
DEFINITION Sequence 1 from Patent WO0214362.
ACCESSION AX467113
VERSION AX467113.1 GI:21900425
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM artificial sequences.
REFERENCE 1
AUTHORS Sallberg, M.
TITLE A hepatitis c virus non-structural ns3/4a fusion gene
JOURNAL Patent: WO 0214362-A 1 21-FEB-2002;
TRIEP AB (SE)
FEATURES
    Location/Qualifiers
    1..2061
        /organism="synthetic construct"
        /mol_type="unassigned DNA"
        /db_xref="taxon:32630"
        /notes="Hepatitis C virus NS3/4A coding region"
ORIGIN
Alignment Scores:
Pred. No.: 1,41e-203 Length: 2061
Score: 3618.00 Matches: 686
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AX467113 (1-2061)
QY      1 MetAlaProIleThrAlaThrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      1 ATGGCCCTATATCGGCCCTATGCCAGACAGACAAAGGGGCTTTGGGATGCATAATCACC 60
QY      21 SerLeuThrGlyArgAspLysAenGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      61 AGCTTGACCGCGCGGACAAAACACAGGTGGAGGTTCAGATCGTGTCAACTGCT 120
QY      41 AlaGlnThrPheLeuAlaThrCysIleAenGlyValCysTrpThrValThrHisGlyAla 60
Db      121 GCCCAGACATTTCTGTGGCAACCTGCATTAACAGGGGTGTGTGGACTGTCTTACCATGGAGCC 180
QY      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

```

```

Db      181 GGAAACAAGGACCAFTTGGTCACTAAGGGTCTCTGTTATCCAGATGTACACCAATGTGGAC 240
QY      81 GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
Db      241 CAAGACCTGTAGCTGGCGCTCCCAAGGTGGCGCTCATTAACACCATGCACACTGC 300
QY      101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db      301 GGCTCTCTGGAGACCTTTACTTGGTCAAGGACCGCCGATGTCTATTCCTTGAAGAGCTCTCTG 360
QY      121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db      361 GGTGATGGCAGGGCAGCTGCTTCCCGCGGCTATCTCTTACTTGAAGAGCTCTCTG 420
QY      141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      421 GGAGGCGCTCTGCTGTGTCGCCCGCAGGACATGCCGTAGGCATATTTCAGAGCGCGGTATGC 480
QY      161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db      481 ACCCGTGGAGTGGCTAAGGCGGTGGACTTCATCCCGTAGAGCTTAGAGACCAACCATG 540
QY      181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db      541 AGGTCCCGGTGTTCTCAGACAACTCTCCACCACGAGTGCCTCCAGAGCTACCAAGTG 600
QY      201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db      601 GCCCACCCTGCATGCTCCACCGCGAGCGGTAAAGACCAAGTCCCGCGCGCATGCACA 660
QY      221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db      661 GCTCAGGGCTACAAGGTGCTGCTCAACCCCTCCGTTGCTGCAACAANTGGGCTTTGGT 720
QY      241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db      721 GCTTACATGTCCAAGGCCCATGGGATTGATCTCAACATCAGGATCGAGTGGGTGAGACAATT 780
QY      261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db      781 ACTACTGGCAGCCCGATCAGTATTCCACCTTACGGCAAGTCTCTTCCGACGGCGGTGT 840
QY      281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db      841 TCAGGGGTGCTTATGACATAATAATTGTGACAGTGCCACTCCACGGATGCAACATCC 900
QY      301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320
Db      901 ATCTTGGGCAATTGGCACTGTCTTGACCAAGCAGAGACCGCGGGGGGAGACTGACTGTG 960
QY      321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      961 CTCGCCACCGGTACCCCTCCGGGCTCGTCACTGTGCCCCCATCTTAACATCGAGAGGTT 1020
QY      341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAATT 1080
QY      361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db      1081 AAGGGGGGAGACATCTCACTCTCTGCACTCAAGAGAGAGGTGCGACGAGCTCGCCGCA 1140
QY      381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      1141 AAACCTGGTTCGGTGGCGCTCAATGCCGTGGCTTACTACCGCGGCTTGATGTGTCTCGTC 1200
QY      401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      1201 ATCCGACCAAGTGTGTGCTGTGCTGGCAACTGACGCCCTCATGACCGGCTTTTACC 1260
QY      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db      1261 GCGGACTTCGATTCGGTGATAGACTGCAACACGTTGTTCACCCAGACAGTCACTTCAGC 1320

```

Qy 441 LeuAspProThrPheThrIleGluThrIleLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTACCTTACCAATGAGCAATCACGCTTCCCGAGGATGCTGCTCCCGTACT 1380
 Qy 461 GluArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaProGly 480
 Db 1381 CAACGTCGGGTAGGACTGCGAGAGGAGCCAGGCACTACAGATTGTGGACCCGGG 1440
 Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuGluGluCysTyrAspAlaGlyCys 500
 Db 1441 GAGCGTCCTTCTGSCATGTTTGACTGCTGCTCTGCGAGTGCTATGACGGGGTGT 1500
 Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaThrMetAenThr 520
 Db 1501 GCTTGGGTATGAGCTTACGCCCGCGAGACCACAGTTAGGCTAGGACATACATGAACACC 1560
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 1561 CCGGCACTTCCCGTGTCCAGACCATCTTGAATTTGGGAGGCGCTTTACGGGTCTC 1620
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIysGlnSerGlyGluAsnLeuProTyr 560
 Db 1621 ACCACATAGACGCCCACTTCCATATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 1681 CTGGTAGCGTACCAAGCACCGTGTGGCTAGAGTTCAGCCCTCCCGCGTGGGAC 1740
 Qy 581 GlnMetTrpIysCysLeuIleArgLeuLeuPheProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGAAGTGTGATCCGCTCTCAAGCCACACCTCCATGGGCAACACCTTGCTA 1800
 Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIysTyrIle 620
 Db 1801 TATAGATGGCGGTGTCCAGATGAAGTCAACCTGACGACCCAGTACCAAGTATATC 1860
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 1861 ATGACATGTATGTCGGTGCACCTGGAGTGTGTCACAGTACCTGGGTGCTGCTGGCGC 1920
 Qy 641 ValLeuAlaIleAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 1921 GTTCTGCTCTGTTGGCCGGTATGCTATCCACAGGCTGCTGTGTATAGTAGTAGG 1980
 Qy 661 IleValLeuSerGlyIysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 1981 ATTGTCTTGTCCGAAAGCCGCAATCATATCCGACAGGGAAGTCTCTACCGGGAGTTC 2040
 Qy 681 AspGluMetGluGluCys 686
 Db 2041 GATGAATGGAAGATGC 2058

RESULT 3

I06434
 LOCUS I06434 5360 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence 487 from Patent EP 0318216.
 ACCESSION I06434
 VERSION I06434.1 GI:590311
 KEYWORDS
 SOURCE Unknown.

ORGANISM

Unknown.

REFERENCE

1 (bases 1 to 5360)
 Houghton, M., Choo, Q.-L. and Kuo, G.
 TITLE Narby diagnostics and vaccines
 JOURNAL Patent: EP 0318216-A1 48-31-MAY-1989
 FEATURES Location/Qualifiers

1..5360
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN

Alignment Scores:
 Pred. No.: 1,74e-200 Length: 5360
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x I06434 (1-5360)
 Qy 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 930 CTGGCGCCCATCAGCGCGTACCCAGCAGACAAAGGGCCCTCTAGGGTGCAATACACC 989
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
 Db 990 AGCCTAATCTGGCCGGGCAAAACCAAGTGGAGGTGAGGTCCAGATTGTCAACTGCT 1049
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 1050 GCCCAACCTTCTCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACACCGGGCC 1109
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1110 GGAACGAGGACCATCGCGTCACCCAGGGTCTCTGCATCCAGATGTATACCAATGTAGAC 1169
 Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1170 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCGCTCATTGACACCTCGCACTTGC 1229
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 Db 1230 GGCTCCTCGGACCTTACCTGGTCACGAGGCACGCCGATGTCATTCCCGTGGCGCGCGG 1289
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 1290 GGTGATAGCAGGGACGCTGCTGTCGCCCGCGGCACGCCGTGGGCATATTAGGGCGGGTGC 1349
 Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 1350 GGGGTGCGCTGTTGTGCCCGCGGCACGCCGTGGGCATATTAGGGCGGGGTGC 1409
 Qy 151 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 1410 ACCCGTGAGTGGCTAAGCGGTGACTTTATCCCTGTGGAGACCTAGAGACACCATG 1469
 Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 1470 AGGTCCCGGTGTTACGGATAACTCTCTCCACCACTAGTGGCCCAAGAGCTTCCAGGTG 1529
 Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 1530 GCTCACTCTCATGCTCCACAGCGAGGCAAAAGCACCAGGTCCCGGTGCTATATGCA 1589
 Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 1590 GCTCAGGCTATAGGTGTAGTACTCAACCCCTCTGTTGTCGCAACACTGGGCTTTGGT 1649
 Qy 241 AlaTyrMetSerLysAlaHisGlyLysProAsnIleArgThrGlyValArgThrIle 260
 Db 1650 GCTTACATGTCACAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACATT 1709
 Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 1710 ACCACTGGACCCCATCATCGTACTCCACTACGGCAAGTTCCTTGGCGAGCGGGTGC 1769
 Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 1770 TCGGGGGCGCTTATGACATAATAATTGTGACAGGTGCCACTCCACGGATGCCACATCC 1829
 Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 1830 ATCTTGGCATCGGCATGTCCTTACCAAGACAGACTGCGGGGGGAGACTGTTGTG 1889


```
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 1890 CTGCCACCCGCCACCCCTCGGCTCGTCACTGTGCCCATCCCAACATCGAGGAGGTT 1949

Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1950 GCTCTGTCCACCCAGGAGATCCCTTTTACGGCAAGGCTATCCCTCCGGAAGTAATC 2009

Qy 361 LysGlyGlyArgHisIlePheCysHisSerIleValCysAspGluLeuAlaIle 380
Db 2010 AAGGGGGAGAGATCTTCATCTTCTGTCAATCAAGAAGAGTGCACCAACTCGCGCA 2069

Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2070 AAGCTGTCTGCATTTGGCATCANTGGCTGCTACTACCGCGTCTTGACGTGTCCGTC 2129

Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2130 ATCCCGACCAAGCGGATGTGTCTGTGGCAACCGATGCCCTCATGCCGCTATACC 2189

Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2190 GGGCATTCGACTCGGTGATAGTACATACGTGTGTACCCAGACAGTCAATTTACAC 2249

Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2250 CTTGACCTACCTTCACTGAGCAATCACGCTCCCGCAGGATGCTGTCTCCGCACT 2309

Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2310 CAACGTGGGGAGAGTGGCAGGGGAGCCAGGATCTACAGATTTGTGGCAGCGGG 2369

Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2370 GAGCGCCCTCGGCATGTTGCACTCGTCCGCTCTGTGTAGTGTATGACGAGGCTGT 2429

Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2430 GCTTGTGTATGAGTCAAGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAAC 2489

Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 2490 CCGGGCTTCCGCTGTGCCAGGACCAATCTGAATTTGGAGGGCTCTTTACAGGCTC 2549

Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2550 ACTCATATAGATCCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACTTCTTAC 2609

Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 2610 CTGGTAGCTACCAAGCCACCGTGTGCGTAGGCTCAAGCCCTCCCACTCGTGAGGAC 2669

Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2670 CAGATGTGAAGTGTGTGCTTCCCTCAAGCCCACTCCATGGGCCAACACCCCTGCTA 2729

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 2730 TACAGACTGGGCGCTGTTCAGATGAATCAACCTGACGACCCCACTCAACAAATACATC 2789

Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 2790 ATGACATGATGTCCGCCACCTCGAGGTGCTACGAGCACCTGGGTGCTCGTGGCGG 2849

Qy 641 ValLeuAlaLeuAlaIleTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 2850 GTCCTGGCTGTGTGGCGGATTTCCCTGTCAACAGGCTGGGTGTCTAGTGGGCGAGG 2909

Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 2910 GTCTGTCTTGTCCGGGAGCGGCAATCATCTGACAGGGAAAGTCTCTTACGAGATTC 2969
```

```
Qy 681 AspGluMetGluGluCys 686
Db 2970 GATGAGTGAAGAGTGC 2987

RESULT 4
LOCUS I09328 5360 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 8 from Patent WO 8904669.
ACCESSION I09328
VERSION I09328.1 GI:587963
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5360)
Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 8 01-JUN-1989;
LOCATION/Qualifiers
1. 5360
source
/organism="unknown"
/mol_type="unassigned DNA"

Alignment Scores:
Pred. No.: 1.74e-200 Length: 5360
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 96.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x I09328 (1-5360)
Qy 1 MetaAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 930 CTGGCGCCCATCACGGCTACGCCAGCAGACAGAGGGGCTCTTAGGGTGCATATCACC 989

Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db 990 AGCCTAATGCGCGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 1049

Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyVal 60
Db 1050 GCCCAACCTTCTGGCAACGTCATCAATGGGTGTGTGACTGTCTACACGGGGCC 1109

Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1110 GGAACGAGGACCATCGGCTCACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 1169

Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1170 CAAGACCTTGTGGGCTGCCCGCTCCCGAAGGTAGCGGCTCATTTGACACCCCTGCACCTGC 1229

Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1230 GGCTCTCGGACCTTACCTGGTCAAGGACGACCGGATGTCATTCCTCGTGGCGGG 1289

Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 1290 GGTGATAGCAGGGGACGCTGCTGCTCCCGGGCCCATTTCTACTTGAAGGGCTCTCTCG 1349

Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1350 GGGGGTCCGCTGTGTGTGCCCGCGGGSCAGCGCGTGGGSCATATTTAGGGCGCGGTGTGC 1409

Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1410 ACCCGTGGAGTGGTAAAGGGGTGGACTTTATCCCTGTGGAGAACCTTAGAGAACCAACATG 1469

Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 1470 AGTCCCGGGTGTTCACGGATATCTCTCCACCGTAGTGTGCCCGGAGACTTCCAGGTG 1529
```

QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 1530 GCTACCTCCATGCTCCACAGGAGCGGAAAGACCAAGTCCCGGCTGCATATGCA 1589
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAunProSerValAlaAlaThrMetGlyPheGly 240
 Db 1590 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTGGT 1649
 QY 241 AlaTyrMetSerLysAlaHisGlyLysLeuAunProSerValAlaAlaThrMetGlyPheGly 260
 Db 1650 GCATTACATGCTCCAGGCTCATGGATCGATCTTCAATCATCAGGACCGGGGTGAGAACAAAT 1709
 QY 261 ThrThrGlySerProThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 Db 1710 ACCACTGGAGCCCACTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCTACCT 1769
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 1770 TCGGGGGGCGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGGATGCCATCC 1829
 QY 301 IleLeuGlyIleGlyThrValLeuAunProSerValAlaAlaThrMetGlyPheGly 320
 Db 1830 ATCTTGGCATCGGCTCTCTTGTGACAGAGAGACTGCGGGGGCGAGACTGGTGTG 1889
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 1890 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGTT 1949
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 1950 GCTCTGTCCACCCGAGAGATCCCTTTTATCGGCAAGCTATCCCTCTCGAGTAAATC 2009
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 2010 AAGCGGGGAGACATCTCATCTTCTGTCAATCAAGAAAGTCCGACCAACTCGCGCA 2069
 QY 381 LysLeuValAlaLeuGlyValAlaAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2070 AAGCTGTGTGATGGGCAATCAATCGCTGGCTTACTTACCGGCTCTGACGTGTCGTC 2129
 QY 401 IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2130 ATCCCGACCGAGCGCGATGTTGCTGCTGGCAACGATGCCCTCATGACCGGCTATACC 2189
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 2190 GCGGACTTCGACTCGGTGATAGTCAATACGTGTGTACCCAGACAGTCGATTTCCAG 2249
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 2250 CTTCACCTTACCTTACCATTTGACAAATCAGCTCCCGGATGCTGTCTCCCGCACT 2309
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 2310 CAACGTCGGGCGAGACTGCGGAGGGGAAAGCCAGGCATCTACAGATTGTGGCACC 2369
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 2370 GAGCGCCCTCCCGCATGTTGCACTCGCTGCTCTGTGAGTGTCTATGACGAGGCTGT 2429
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuAlaTyrMetAsnThr 520
 Db 2430 GCTTGGTATGACTCAGCGCCCGCGAGACTACAGTTAGGTACGAGCGGTACATGAACACC 2489
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
 Db 2490 CCGGGGCTTCCCGGTCCAGGACCACTTGAATTTTGGAGGGGCTCTTTACAGGCTC 2549
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 2550 ACTCATATAGATGCCACTTCTTATCCAGCAAGCAGAGTGGGAGAACCTTCTCTTAC 2609
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580

Db 2610 CTGGTAGCGTACCAAGCCACCGTGTGGCTAGGCTCAAGCCCTCCCTCCATCGTGGAC 2669
 QY 581 GlnMetTpyLysCysLeuIleAsgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 2670 CAGATGTGGAGTGTGTTGATTCGCTTCAAGCCACCTCTCATGGGCCCAACACCCCTGCTA 2729
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 2730 TACAGACTGGCGCTGTTCAGATGAATCACCCTCAGCCACCCAGTCACCAATATCATC 2789
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 2790 ATGACATGTCATGTGGCGGACCTGGAGTGTCTCAGGACACCTGGGTCTCTTGGCGGC 2849
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 2850 GTCCCTGGCTGTTTGGCGGCTATTGCTGTCAACAGCTGCGTGTCTAGTGGGCGAG 2909
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 2910 GTCGTCTTGTCCGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 2969
 QY 681 AspGluMetGluGluCys 686
 Db 2970 GATGAGATGGAGAGTGC 2987

RESULT 5
 I06440
 LOCUS I06440 6785 bp DNA linear PAT 02-DEC-1994
 DEFINITION Sequence:54-from Patent EP 0318216.
 ACCESSION I06440
 VERSION I06440.1 GI:590312
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 6785)
 AUTHORS Houghton, M., Choo, Q.-L. and Kuo, G.
 TITLE Nanv diagnostics and vaccines
 JOURNAL Patent: EP 0318216-A1 54 31-MAY-1989;
 FEATURES
 Location/Qualifiers
 1..6785
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Alignment Scores:
 Pred. No.: 2,32e-200 Length: 6785
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x I06440 (1-6785)
 QY 1 MetaAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleThr 20
 Db 1203 CTGGCGCCCATCAGCGGTACGCCAGCAGCAAGGGGCTCTCTAGGTGTCATTAATCACC 1262
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 1263 AGCTTAACCTGCGCGGACAAAACCAAGTGAGGGTGGGTCAGATTGTGTCACTGCT 1322
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 1323 GCCCAAACTTCTTGGCAACGTGTCATCGGTGTGTCTGACTGTCTACACGGGCGC 1382
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValasp 80
 Db 1383 GGAACGAGGACCATCGCTCGTCAACCAAGGGTCTCTGTCTATCCAGATGTATACCAATGTAGAC 1442

Qy 81 GlnAspLeuValGlyThrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1443 CAAGACCTTGTGGCTGGCGCTCCGCAAGAGTAGCGCTCATTCACACCCCTGCACTTGC 1502
Qy 101 GlySerSerAspLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 1503 GGCTCCTCGACCTTACCTTGTGTACGAGGACGCGGATGCTATTCCCGTGGCCGCGG 1562
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuYusGlySerSer 140
Db 1563 GGTGATAGAGGGGACGCTGCTGTGCGCCCGGCCCATTTTCTTACCTTGAAGGCTCCTCG 1622
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1623 GGGGTCCTGCTGTGTGCGCGCGGACGCGCTGGCATATTTAGGCGCGGGTGTGC 1682
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1683 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACACCATG 1742
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 1743 AGGTCCCGGTGTTCAGGATACTCTCTCCACCAAGTAGTGCCCGCAGAGCTTCCAGGTG 1802
Qy 201 AlaHisLeuHisAlaProThrGlySerGlySerThrLysValProAlaAlaValAla 220
Db 1803 GCTCACCTCCATGCTCCACAGGACGGGCAAGCACCAAGTCCCGGCTGCATATGCA 1862
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 1863 GCTCAGGGCTATAGGTGTGTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 1922
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1923 GCTTACATGTCCAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAT 1982
Qy 261 ThrThrGlySerProIleThrThrThrThrThrThrThrThrThrThrThrThrThr 280
Db 1983 ACCACTGGACGCCCATCATGCTACTCCACTACGGCAAGTTCCTTCCCGACGGCGGTGC 2042
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2043 TCGGGGGCGCTTATGACATAATAATTGTGACGAGTGCCACATCCACGATGCCATCC 2102
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2103 ATCTTGGCATCGCACTGCTTGTACCAAGCAGAGACTCGGGGCGAGACTGGTTGTG 2162
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2163 CTCGCCACCGCCACCTCGGCTCGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2222
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2223 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAGTAATC 2282
Qy 361 LysGlyGlyArgHisIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2283 AAGGGGGGAGACATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2342
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2343 AAGCTGTCTGCATTTGGCAATCGGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2402
Qy 401 IleProThrSerGlyAspValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2403 ATCCGACGACGCGGATGT 2462
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2463 GCGGACTTCACCTCGGTGTAGACTGCAATACGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2522
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460

Db 2523 CTTGACCCCTACCTTCAACATTGAGACAATCACTCCGCTCCCGCAGGATGCTGTCTCCGCACT 2582
Qy 461 GlnArgArgGlyValThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2583 CAACGTCGGGCGAGACATGCGAGGGGAGGCGAGCACTTACAGATTTGTGGCACCGGG 2642
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2643 GAGCGCCCTCCGCGCATGTCGACTCGTCCGCTCTCTGTGAGTGTATGACGCGAGCTGT 2702
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2703 GCTTGTATGAGCTCACCCCGCAGACTACAGTTAGCTACGAGCGGTACATGAACACC 2762
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGlyValPheThrGlyLeu 540
Db 2763 CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCCTC 2822
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2823 ACTCATATAGATGCGCCACTTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCCTTAC 2882
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 2883 CTGGTAGCTACCAAGCCACCGTGTGCTAGGCTCAGGCCCTCCCGCATCGTGGGAC 2942
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2943 CAGATGTGGAAGTGTGTTGATTCGCTCTCAAGCCCACTCCATGGGCCCAACACCCCTGCTA 3002
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysThrIle 620
Db 3003 TACAGACTGGCGCTGTTCAGATGAATCACCCTGACGCCACCCAGTCCCAAAATACATC 3062
Qy 621 MetThrCysMetSerAlaAspLeuValValThrSerThrThrThrThrThrThrThrThr 640
Db 3063 ATGACATGATGTCCGCGACCTGGAGGTGCTCAGAGACCTGGGTGCTCGTGGCGGC 3122
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3123 GTCTTGGCTGCTTTGGCGCGCTATTGCTGTCAACAGGCTCGTGGTCTATAGTGGGCGG 3182
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3183 GTCGCTTGTTCGGGAGCCGCGCATCATCTCAGAGGAAGTCCCTTACCGAGAGTTC 3242
Qy 681 AspGluMetGluGluCys 686
Db 3243 CATGAGATGGAAGATGC 3260
RESULT 6
I09329
LOCUS I09329 6785 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence (10) from Patent WO 8904669.
ACCESSION I09329
VERSION I09329.1 GI:587964
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 6785)
AUTHORS Houghton M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 10 01-JUN-1989;
FEATURES
Location/Qualifiers
1..6785
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores: 2.32e-200 Length: 6785
Pred. No.: 3574.00 Matches: 672
Score:

Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	6	Gaps:	0
US-09-930-591-2 (1-686) x 109329 (1-6785)			
Qy	1	MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
Db	1203	CTGCGGCCCATCACGGGTACGCCAGCAGACAGAGGGGCTCTAGGGTGCAATCAC	1262
Qy	21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla	40
Db	1263	AGCCTAACTGGCGGGCAAAACCAAGTGGAGGGTGAGGTCCAGATTGTCTCAACTGTT	1322
Qy	41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIleThrValTyrHisGlyAla	60
Db	1323	GCCCAAACTTCCTGGCAACGTGCATCAATGGGGTGTGCTGGACTGTCTACCAAGGGGCC	1382
Qy	61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
Db	1383	GGAAACGAGGACCATCGGCTCACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC	1442
Qy	81	GlnAspLeuValGlyTyrProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys	100
Db	1443	CAAGACCTTGTGGCTGGCCGCTCCCAAGGTAGCGCTCATTTGACACCCCTGCATTGGC	1502
Qy	101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIlePheArgAlaValCys	120
Db	1503	GGCTCTCGGACCTTACCTGGTCACGAGCAGCGCGATGTCATTCGCGTGGCCGCGG	1562
Qy	121	GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
Db	1563	GGTGATAGAGGGGACGCTGCTGCTGCCCGCCGCTTCTTCTTCTTGAAGGCTCTCTCG	1622
Qy	141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaValCys	160
Db	1623	GGGGTCCGCTGTGTGCCCCGGGGCACCGCTGGGCATATTTAGGGCCGGGTGTGC	1682
Qy	161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
Db	1683	ACCCGTGGAGTGGTAAAGCGGTGGACTTATCCCTGTGGAGAACCTTAGAGCAACCATG	1742
Qy	181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
Db	1743	AGTTCCTCGGTGTTCAGGGATACCTCTCTCCACAGTAGTGGCCCCAGAGCTTCCAGG	1802
Qy	201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
Db	1803	GCTCACCTCCATGCTCCACAGGCGAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA	1862
Qy	221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
Db	1863	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTCTGCTGCAACACTGGGCTTTGGT	1922
Qy	241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
Db	1923	GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT	1982
Qy	261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
Db	1983	ACCACTGGAGCCCATCAGACTCTCCACCTACGGCAAGTTCCTTGGCGAGCGGGGTGC	2042
Qy	281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
Db	2043	TCGGGGGGCGCTTATGACATAATAATTTGTGACAGTGGCCACTCCACGGATGCCACATCC	2102
Qy	301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
Db	2103	ATCTTGGGCACTCGGCACTGCTCTTGACCAAGACAGACTCGGGGGGAGACTGGTGTG	2162
Qy	321	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal	340
Db	2163	CTCGCCACCGCCACCCCTCCGGGCTCCGTCTACTGTGCCCATCCCAACATCGAGGAGTT	2222
Qy	341	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
Db	2223	GCTCTGTCCACCCGCGAGAGATCCCTTTTACGCAAGGCTATCCCTCCGAAGTAATC	2282
Qy	361	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
Db	2283	AAGGGGGGACACATCTCATCTTCTGTCTTCAAGAGAGAGTCCGACCACTCGCCGCA	2342
Qy	381	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
Db	2343	AAGCTGTGCTGATGGGCATCAATCGGTGGCTTACTACCGCGTCTTGAGTGTCCGTC	2402
Qy	401	IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr	420
Db	2403	ATCCGACAGCGGCGATGTTGTCTGTGGCAACCGATGCCCTCATGACGGGTATACC	2462
Qy	421	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
Db	2463	GGCGACTTCGACTCGGTGATAGACTGCAATACGHTGTACCCAGACAGTCGATTCAGC	2522
Qy	441	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
Db	2523	CTTGACCTCATCTTCAACATTGAGCAATCAGCTCCCGAGGATGCTCTCCCGCACT	2582
Qy	461	GlnArgArgGlyArgThrGlyArgGlyArgGlyIleTyrArgPheValAlaProGly	480
Db	2583	CAAGTCGGGGCAGGACTGGCAGGGGGAAGCCAGGATCTACAGATTTGTGGCACCGGG	2642
Qy	481	GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys	500
Db	2643	GAGCGCCCTCCGCGATGTTCTGACTCTGCTCGTCTCTGTCAGTGTATGACGAGGCTGT	2702
Qy	501	AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr	520
Db	2703	GCTTGTATGAGCTACGCGCCGCGAGACTACAGTTAGGCTACAGGCTACATGAACACC	2762
Qy	521	ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu	540
Db	2763	CCGGGGCTTCCGTGTGCCAGGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCTC	2822
Qy	541	ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr	560
Db	2823	ACTCATATAGATGCCCACTTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC	2882
Qy	561	LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp	580
Db	2883	CTGGTAGGCTACCAAGCCCGCTGTGCTAGGCTCAAGCCCTCCGCCATCTGGGAC	2942
Qy	581	GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu	600
Db	2943	CAGATGTGGAAGTGTGTTGATTCCGCTCAAGCCCACTCCATGGGCCAACACCCCTGCTA	3002
Qy	601	TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle	620
Db	3003	TACAGACTGGGCGCTGTTTCAAGATGAATACCTCCCTGACGACCCAGTCACCAATACATC	3062
Qy	621	MetThrCysMetSerAlaAspLeuGluValValThrSerThrThrValLeuValGlyGly	640
Db	3063	ATGACATCATGTCGGCGGACCTGGAGTCTGCACGACACCTGGGTCTGCTGGGGC	3122
Qy	641	ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg	660
Db	3123	GTCTCTGGCTGTTTGGCCGCTATTGCTGTCAACAGCTGCGGTGCATAGTGGGAGG	3182
Qy	661	IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe	680
Db	3183	GTCTCTTGTCCGGGAGCCCGCAATCATCTTACAGGGAAGTCTCTTACCGAGAGTTC	3242
Qy	681	AspGluMetGluGluCys	686
Db	3243	GATGATGGAAGAGTGC	3260

```

RESULT 17
AR118696 LOCUS AR118696 7310 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence (74) from patent US-6150087.1
ACCESSION AR118696
VERSION AR118696.1 GI:14100606
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1--(bases 1 to 7310)
  Chien D.Y.
  TITLE NANBV diagnostics and vaccines
  JOURNAL Patent: US 6150087-A 74 21-NOV-2000;
  FEATURES
    Location/Qualifiers
      1..7310
        /organism="unknown"
        /mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,54e-200 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AR118696 (1-7310)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1728 CTGGCGCCATACCGGCTACGCCGACACACAGGGCCCTCTAGGTGCATAATCAC 1787
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
DB 1788 AGCTAACTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGCACTGCT 1847
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 1848 GCCCAAACCTCTCTGGCAACGTGCATCAATGGGGTGTGGACTGTCTACCAAGGGGCC 1907
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 1908 GGAACGAGGACCATCGGTCACCCAGGGTCCGTGCATCCAGATGTATACCAATGTAGAC 1967
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 1968 CAAGAQTGTGGGCTGGCCCGCTCGCAAGGTAGCGGCTCATTGACACCCCTGACTTGC 2027
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 2028 GGCTCTCGACCTTACCTTGGTACAGAGCAGCCGATGTCAITCCCGTGGCCGGCGG 2087
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 2088 GGTGATAGCAGGGGACGCTGTGTGCCCCCGGCCCATTTCTTACTTTGAAAGGCTCCTCG 2147
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaIleValCys 160
DB 2148 GGGGGTCCGCTGTGTGCCCGGGGACGCCGTGGGCATATTTAGGGCCGGGTGTC 2207
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 2208 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGCAACCAUG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 2268 AGGTCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGCACAGAGCTTCCAGGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220

```

```

2328 GCTCACTCCATGTCTCCACAGGCGCGCAAAAGCCACCAAGGTCCCGGCTGCATATGCA 2387
221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
2388 GCTCAGGGCTATAAGTCTAGTACTCAACCCCTCTGTGTCTGCAACACACTGGGCTTGGT 2447
241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
2448 GCTTACATGTCAAAGGCTCATGGATCGATCCTAACATCAGACCGGGGTGAGAACAAAT 2507
261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
2508 ACCACTGGCAGCCCATCACTACTCCACTACGGCAAGTTCCTTCCGACGGCGGTGC 2567
281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
2568 TCGGGGGCGGTTATGACATAATAATTTGTACGAGTGCCACTCCACGGATGCCACATCC 2627
301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
2628 ATCTTGGGCATCGGCACCTGTCTTGACACAGAGACTGGGGGGCGAGCTGGTTGTG 2687
321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal 340
2688 CTCGCCACCGCACCCCTCCGGGCTCCGTCCTCACTGTGCCCATCCCAACATCGAGAGGT 2747
341 AlateuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
2748 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGTATCCCCCTCGAAGTAATC 2807
361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
2808 AAGGGGGGAGACATCTCATCTCTGTCTTCAAAGAAGAAGTGCAGCAACTCGCCGCA 2867
381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
2868 AAGCTGTGTGCGATTGGGCATCAATGCCGTGCGCTACTACCGGCTGTGAGCTGTCCGTC 2927
401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
2928 ATCCGACAGCGGGGAGTGTGTCTGTGGCAACCGATGCCCTCATGACCGGCTATACC 2987
421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
2988 GCGCATCTCGACTCGTGATAGACTGCAATACGTGTGTCAACACAGACAGTGCATTTACG 3047
441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
3048 CTGTACCTTACCTTACCATTGAGCAATACGCTCCCGCCAGGATGCTGTCTCCCGCAT 3107
461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
3108 CAAAGTCCGGGCGAGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGG 3167
481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
3168 GAGCCCGCTCCCGCATGTTGACTTCGCTCGCTCTGTGAGTGTCTATGACGAGGCTGT 3227
501 AlaTyrProGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
3228 GCTTGGTATGAGCTACCGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3287
521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
3288 CCGGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCCCTC 3347
541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
3348 ACTCAATAGATGCCCATTTCTATCCCGACAAAGACAGAGTGGGAGAACTTCTCTTAC 3407
561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
3408 CTGGTAGGCTACCAAGCCACCGTGTGCGTCTAGGGGCTCAAGCCCTCCCCCATCGTGGAC 3467

```

Qy 581 GlnMetTrpLysCysLeuIleAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrlle 600
Db 3468 CAGATGTGGAAGTGTGTTGATTGGCTCAAGCCACACCCCTCCATGGCGCAACACCCCTGCTA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrlle 620
Db 3528 TACAGATGGCGCTGTTCAGATGAATCAACCTGACGACCCAGTCAACCAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGATGTCGGCGACCTGGAGTGTGTCACGAGACCTGGTGTGTCGTCGCGGC 3647
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValLeuValGlyArg 660
Db 3648 GTCTGTGCTGCTTTGGCGCGGTATGCTGTCAACAGCGCTGCGTGTGTCATAGTGGGCAGG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleProAlaIleProAspArgGluValLeuValTrpValGluPhe 680
Db 3708 GTGCTGTGTCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGATGC 3785
RESULT 8
LOCUS I09331 7310 bp DNA linear PAT 02-DEC-1994
DEFINITION Sequence 15 from Patent WO 8904669.
ACCESSION I09331
VERSION I09331.1 GI:587966
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7310)
AUTHORS Houghton, M., Choo, Q.-K. and Kuo, G.
JOURNAL Patent: WO 8904669-A 15 01-JUN-1989;
FEATURES
Location/Qualifiers
1..7310
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2.54e-200 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
US-09-930-591-2 (1-686) x I09331 (1-7310)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGGCCCATACACGGGTACGCCACGACAGGGGGCCCTCTAGGGTGCAATACACC 1787
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
Db 1788 AGCCTAACTGGCGGGACAAACCAAGTGGAGGTGAGGTGTCAGATTGTCAACTGCT 1847
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1848 GCCCAACCTTCCTGGCAACGTGATCAATGGGGTGTGCTGATCCAGATGATACCATGTAGAC 1907
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGAACGAGGACCATCGGTACCCAGGGTCTGTCTCATGATGATACCATGTAGAC 1967
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAGACCTTGTGGGCTGGCGGCTCCGCAAGGTAGCGGCTCATTGACACCCCTGCACTTGC 2027

Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgArg 120
Db 2028 GGCCTCTCGGACCTTTACCTGGTCACGAGGACCCGATGTCAATTCGCGTCGCGCGCGG 2087
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 2088 GGTGATACAGGGGGCAGCTGCTGTCGCCCCGGGCCATTTCTACTTGAAGGCTCTCG 2147
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 2148 GGGGGTCCGCTGTTGTGTCGCGCGGGCACGCCGTGGCATATTTAGGGCGCGCTGTGC 2207
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
Db 2208 ACCCGTGGAGTGGCTTAAGCGGTGACTTTATCCCTGTGGAGAACCTAGAGACAACCATG 2267
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCGGTGTTCAAGGATAACTCTCTCCACCACTAGTGTCCCGAGCTTCCAGGTG 2327
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GCTCACCTCCATGCTCCACAGGCGGCAAGAGCAACAGGTCCCGGTGCAATATGCA 2387
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 2447
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 2448 GCTTACATGTCCAAAGGCTCATGGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAAT 2507
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCCATCACGTACTCCACTACGCGAAGTCTCTTGGCGCGGGGTGC 2567
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCCAGGATGCCACATCC 2627
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTTGGCATCGGCATGCTCTTGACCAAGCAGAGACTCGCGGGGGGAGACTGTTGTG 2687
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTGCGCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGAGGTT 2747
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGlyAlaIle 360
Db 2748 GCTCTGTCCACCGGAGAGATCCCTTTTACGCGAAGGCTATCCCCCTCGAAGTAAATC 2807
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCTCATCTTCTGATCAAGAAGAGTCCGACGAACTCCCGCGCA 2867
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGTGTGTCGATGGGCATCAATCGCGTGGCTTACCGCGGTCTTGACGTCTCGCTC 2927
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCGACCGCGCGATGTTGTGTGTGGCAACCATGCCCTCATGACCGGTATACC 2987
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGAGTCTTTCAGC 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCTTCACTTCACTTGAACAATCAGCTCCCCAGGATGCTGTCTCCCGACT 3107


```
Qy 461 GlnArgArgGlyThrGlyArgGlyLysProGlyLysLeuValAlaProGly 480
Db 3108 CAACGTCGGGCGAGGACTGCAGCGGGGAGCCAGGCACTTACAGATTGTGGCAGCGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysValAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCGGCATGTTGCACTCGTCCGTCCTCTGTGAGTGCTATGACGAGAGCTGT 3227
Qy 501 AlaTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGGTATGAGTTCAGCGCCGCGAGACTACAGTATAGGCTACGAGCGTACATGAACAC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTCCCGTGTGCCAGGACCATCTTGAATTTGGAGGCGCTCTTACAGCGCTC 3347
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCCTTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 3408 CTGGTAGCGTACCAAGCCCGTGTGGCTAGGCGCTCAAGCCCTCCCATCGTGGGAC 3467
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCAAGCCCTCCATGGGCCAACCCCTGTCTA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGCGCGTGTTCAGATGAATCAACCTCAGCGCACCCATGACCAAAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 3588 ATGACATGATGTGCGCGACCTGGAGGTGCTCAGCAGACCTGGTGTGTCGGCGC 3647
Qy 641 ValLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValleValGlyArg 660
Db 3648 GTCTGCTGCTTTGGCGCGTATTGCTGCTCAACAGGCTGCTGCTAGTGGCGAG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTGCTGTGTCGGGAGCGGCAATACCTACGAGGGAAGTCTCTACCGAGAGTTC 3767
Qy 681 AspGluMetGluGlyCys 686
Db 3768 GATGAGTGAAGAGTGC 3785

RESULT 9
HPCPOLYP
LOCUS
DEFINITION
Hepatitis C virus polyprotein gene, partial cds.
ACCESSION
M32084
VERSION
M32084.1
GI:329875
KEYWORDS
polyprotein.
SOURCE
Hepatitis C virus
ORGANISM
Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE
1 (bases 1 to 7310)
AUTHORS
Choo, Q.-L., Richman, K. and Han, J.
TITLE
The nucleotide sequence of the Hepatitis C viral genome
JOURNAL
Unpublished (1990)
COMMENT
Original source text: Hepatitis C virus, cDNA to viral RNA, clones
K9-1 through 15e, isolated from chimpanzee (individual 910) blood
plasma.
Draft entry and printed sequence for [1] kindly submitted by
M'Houghton, 22-FEB-1990; Chiron Corporation, 4560 Horton Street,
Emeryville CA 94608.
FEATURES
Location/Qualifiers
1..7310
/organism="Hepatitis C virus"
/mol_type="genomic RNA"
```

CDS

```
/db_xref="taxon:11103"
<1..7310
/feature="polyprotein"
/codon_start=3
/protein_id="AAA45677.1"
/db_xref="GI:329876"
/translation="GCPRLASCPPLTDFDQGGPI SVANGSGDQBPYCWHPKPC
GIVFAKSCGVFYCFYTFSPVVGTTDRSGAFYKSGENDTDVFNINRTRPLGNWFGC
TWMSNTGFTKVGCGAPPVGGAGNNTLHCPDCKRPKPDATGSCGGSPWITPLCLVD
YFYLWHPYCTINYTIKIRMYGVGVEHLPEACNACNTRGDCRDLSDRSLSLLLT
TQOVLPCSTFTLSTALSTGLIHQNIYDQVLYGVGSSIASWAKWEYVLLFLL
ADAVSCSLMMMLLIISOAALENLVINAASLAGTHGLSVLFYFCFAYLKGKWP
GAVTFYGMPLLLALLPORVALDEVAASCGVVLGLMALTLSPYKRYISWC
LWVYFLYRVEAOLHWIPIPLNVRGGRDAVILAMCAVHPTLFDITKLILAGPLW
ILOSLKVPYFVRVQGLRRCALARMIGSHYQMVIIKLCALTEGYVYNHILPLRD
WAHNGRLDAVAVPVSQMETKLTITWGAADTAACGDIINGLPVSARGRILLUGPAD
GMVSGKRLAPITAYAQOTRGLGCIITSLTRDKNQGEVEGIVTAAGTFLATCI
NGVWTYVHGAGRTIATSPKPGIYOMYTNVDQLVGWPAPQSGRSLLTFCGSDLYL
VTRADVLPVRRRGDSRGLSPPI SYLKGSGGGLICPAHAGVGI FRAAVCTRGVA
KAVDPI PVENLETTMRSPVFTDNSPPVVPQSFQVHLHAPTGSKGKTKVFAAAGG
KVLVLPASVAATLGFAGYMKAGIDENIETGVTITGSPITYSTYKGLADGGC
GGADIIICDECHSDATSLIGITVLDOAETAGARLVLATATPPGVTVPHNIEE
VALSTTGIPFYKAIPLIEVINGRHLIFCHSKKCDLAAKLVAGINAVAYRGLD
VSVPTSGDVVVAIDALMTGTDGDFSDIDCNCVTQTVDLSLDTFTIITLPQD
AVSTQRRGTGRKPGIYFVAPCRPSGFMFDSVLCECYDAGCAYELTAPETVR
LAYNNTPLPVCODHLEFWEGVFTGLTHIDAEFLOTQKSGENLYLVATQATVCR
ACAPPSWDQMKGLIRLPTLHGPTPLLYRGLAVONEITLTPVTKYIMTCMSADLE
VVTSTVLVGGVLAALAAKYLSTGCVNVGVSLGKPAIIPDREVLYRBEDEECS
QHLPIEGOMLAEQFKQKALGLOTSRQREVIAPAVQTNWQKLETFWAKHMPFIS
GIOLAGLSTLPGNPALASLMAFTAAVTSPLTTSOTLTLNGLGWAQAQLAAPGAATA
FVGLAGAAATGSGVLGKVLIDILAGYAGVAGALVAFKMSGEVPTEDLVNLLPAI
LSPGALVVGVTQAAAILRRHVGGEGAVQMNRLIAFASRGNHVSPTHYVPSDAAARV
TALSLSLTVOLLRLRHLQWISSECTPCSGSLRDIWDWICEVLSDFKTLKAKLMPQ
LPGIPFVSCORGKGVMEVDGIMHTRCHGAEITGHVKNGTMRIVGPTCRNMSGTF
PINATYTGCTPLPAPNYTFALWYSAEYVEIRQVGDGHVYVTGMTDNLKCPQVPS
PEFTLGLDRLHRFPAPPCALLREVSFRVGLHEYFVGSQPCPEPFDVAVLISMLT
DPHITABAGRIARLARGSPSVASSASQSLASLKAATCTANHDSPOAELLEANLWR
CEMGNITRVESNWLIDSFPLVAEDEEISVPABILKHSRRAQALFWARPD
YNPLVETWGTDEYPPVPHGCPPLPPKPPPPPPKRTVLTTESTLSTALALATR
VSPGSSVTGIGTNDNTSSBPAGCPDSDAESVSNMPLEGGPGDPLDGSWSVTR
SSEANEDVCCMSYSWTGALVTPCAAEOKLPINALSNLSLRHNLVYTTSSAC
QRQXVDFRLQVLDSHYQDVLKEVKAASKVXANLLSVEACSLTFPHSKSKFVGC
AKDVRCHAKRATHINSVQKLDENVPTITIMAKNEVFCVQPEKGRPKALIVF
PDLGVRCENKALYDVVTKLPLVNGSSYGFQSPQRVEFLVQAKSKKTPMGFSYD
TRCFDSTVTESDIRTEALYQCDLDPQARVAIKSLTERLYVGGFLNRSNGCYRR
CRASGLTSCGNTLTCVTKARAAACRAAGQDCTMLVCGDLYVICSAGVQEDAAAL
RAFTAMTRYAPPPDPOPEYDLELITSCSNVSVADHDAGKRVYLTRODPTPLAR
AAWETARHTPVNSMLNLIIMFAPTLWARMILMTHFFSVLIARDQLEQALDCEIYGACY
SIEPLDLPPIIQL"
```

ORIGIN

```
Alignment Scores:
Pred. No.: 2,548-200 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 14 Gaps: 0

US-09-930-591-2 (1-686) x HPCPOLYP (1-7310)
Qy 1 MetaProteinTherAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGCGCCCATCAGCGGTGACCCAGCAGACAGGGGCCCTCTTAGGTGTCATATACC 1787
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThra 40
Db 1788 AGCCTAACTGGCGCGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCACTGT 1847
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCystTrpThrValTyrHisGlyAla 60
Db 1848 GCCCAACACCTTCCTGGCAACGTGCATCAATCGGGGTGTCTGTGACTGTCTACACCGGGCC 1907
```


QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1908 GGACGAGGACCATCGCGTACACCAAGGTCCTGTCTACATCCAGATGATATACCAATGTAGAC 1967
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1968 CAAGACCTTGTGGCTGGCGCGTCCGCAAGGTAGAGCGCTCATTTGACACACCTGCATTCG 2027
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValAlaArgArg 120
 Db 2028 GGCTTCCTCGACCTTTACCTGGTCAAGGACCGCCGATGTCACTCCGCTGGCCGGCGG 2087
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 2088 GGTGATAGACGGGCGAGCTGCTGCTGCGCGCGCCATTTCTACTTGAAGGCTCTCTCG 2147
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 2148 GGGGCTCGCTGTGTGCGCGCGGGGACCGCGTGGGCATATTTAGGGCGCGGTGTGC 2207
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGlnSerLeuGluThrThrMet 180
 Db 2208 ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCTCTCGGAGAACCTTAGAGACAACCATG 2267
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 2268 AGGTCCCGCGGTTCACGGATAACTCTCTCTCCACAGTAGTGGCCCCAGAGCTTCCAGGTG 2327
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 2328 GCTCACTCCATGCTCCACAGGCGAGCGGCAAGACCAAGGTCCCGGTGCATATGCA 2387
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 2388 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTCTTCTGCAACACCTGGGCTTTGGT 2447
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaArgThrIle 260
 Db 2448 GCTTACATGCTCAAGGCTCATGGATGATGATCACTAGGACCGGGGTGAGACAAAT 2507
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 2508 ACCACTGGCAGCCCATCACGTACTCACCTACGGCAAGTCTCTTGGCGGCGGCGGTGC 2567
 QY 281 SerGlyGlyAlaTyrAspIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 300
 Db 2568 TCGGGGGCGCTTATGACATATATTTGTGACGAGTGCCACTCCCGATGCCACATCC 2627
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2628 ATCTTGGGCATCGGCACCTGCTTGACCAAGCAGAGACTGCGGGGGGAGACTGGTTGTG 2687
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 2688 CTCGCCACCGCCCTCCCGGCTCGCTCACTGCTGCCCATCCCAACATCGAGGAGGTT 2747
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2748 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCGCTCGAGTAATC 2807
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 Db 2808 AAGGGGGGAGACATCTCATCTTCTGTCAITCAAGAGAGAGTGGCAGAACCTCGCGCA 2867
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2868 AAGCTGTGCGATTTGGGCGATCAATCGCGTCTACCGGCTCTTACGCTGCTCGCTC 2927
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2928 ATCCGACACGCGCGATGTTGTCTGCTGGCAACCGATGCCCTCATGACCGGCTATACC 2987
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

Db 2988 GGCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACCCAGACAGTTCGATTCAGC 3047
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 3048 CTTGACCTTACCTTCCACATTCAGACAATACGCTCCCGCAGGATGCTGTCTCCGCACT 3107
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 3108 CAACGTCGGGGCAGGACTGGCAGGGGGAAGCAGGCACTTACAGATTTGTGGCACCGGG 3167
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAsnAlaGlyCys 500
 Db 3168 GAGCGCCCTTCGGCATTTGACCTGCTCCGCTCTCTGTAGTGTATGACGCGAGGCTGT 3227
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 3228 GTTGTGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3287
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
 Db 3288 CGGGGCTTCCGCTGTGCGAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCTC 3347
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 3348 ACTCATATAGATGCCACTTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCTTAC 3407
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerThrAsp 580
 Db 3408 CTGTGTAGCTTACCAAGCCACCGTGTGCTAGGCTCAAGCCCTCCCGCATCTGGGAC 3467
 QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 3468 CAGATGTGGAGTGTGTGATTCGCTCAAGCCCGCTCCATGGGCCACACCCCTGCTA 3527
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 3528 TACAGACTGGGCGCTGTTGAGATGAAATCACCTGACGACCCAGCAGTACCAATACATC 3587
 QY 621 MetThrCysMetSerAlaAspLeuValValThrSerThrThrTrpValLeuValGlyGly 640
 Db 3588 ATGACATGCATGTGGCGGACCTGGAGTGTGTCAGGACCTGGGTGCTGTTGGCGGC 3647
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 3648 GTCTGTGCTGCTTGGCGCGTGTGCTGCTGCAACAGGCTCGTGGTGCATAGTGGCAGG 3707
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 3708 GTCTGTCTTGTCCGGGAGCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 3767
 QY 681 AspGluMetGluGluCys 686
 Db 3768 GATGAGATGGAGAGTGC 3785

RESULT 10
 I08294
 LOCUS
 DEFINITION Sequence 1 from Patent-EP 0388232
 ACCESSION I08294
 VERSION I08294.1 GI:588994
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 9185)
 AUTHORS Houghton, M., Choo, Q.-L., and Kuo, G.
 TITLE NANV diagnostics and vaccines
 JOURNAL Patent, EP 0388232-Al-1-19-SEP-1990;
 FEATURES
 1..9185 Location/Qualifiers
 /organism="unknown"
 /mol_type="unassigned DNA"

Db 5375 GTCTCTTGTCGGGAAGCGGCAATCATACCTGACAGGAAGTCTCTTACCGAGAGTTC 5434

Qy 681 AspGluMetGluGluCys 686

Db 5435 GATGAGATGGAAGAGTGC 5452

RESULT 11

LOCUS BD091382 9185 bp DNA linear PAT 27-AUG-2002

DEFINITION HCV cultivation method in eucaryotic cells.

ACCESSION BD091382

VERSION BD091382.1 GI:22636993

KEYWORDS JP 2001314192-A/3.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 9185)

AUTHORS Weiner, A.J., Steimer, K.S. and Houghton, M.

TITLE HCV cultivation method in eucaryotic cells

JOURNAL Patent: JP 2001314192-A 3 13-NOV-2001;

CHIRON CORP

COMMENT OS Homo sapiens (human)

PN JP 2001314192-A/3

PD 13-NOV-2001

PF 15-MAR-2001 JP 2001075114

PR 25-AUG-1989 US 398667

PI AMY J WEINER, KATHELYN S STEIMER, MICHAEL HOUGHTON PC

C12N15/09, C12N5/10, C12N7/00, C12R1/93, C12N15/00, PC

C12N5/00

CC HCV cultivation method in eucaryotic cells

FH Key Location/Qualifiers

FT source 1..9185

FT /organism='Homo sapiens (human)'.
Location/Qualifiers
1..9185
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'

FEATURES

source

ORIGIN

Alignment Scores:

Pred. No.: 3,36e-200 Length: 9185

Score: 3574.00 Matches: 672

Percent Similarity: 99.42% Conservatives: 10

Best Local Similarity: 97.96% Mismatches: 4

Query Match: 98.78% Indels: 0

DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x BD091382 (1-9185)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20

Db 3395 CTGGCGCCCATCACGGCGTACGCCAGACAGACAGGGGCGCTCTAGGGTGCAATCAACC 3454

Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40

Db 3455 AGCTTACTGCGCGGCAAAACCAACAGTGGAGGTGAGGTCCAGATTGTGTCACTGCT 3514

Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60

Db 3515 GCCCAACCTTCCTGGCAACGTGATCAATGGGGTGTGCTGGAGTGTCTACCAACGGGGCC 3574

Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

Db 3575 GGAACGAGGACATCGGTACCCCAAGGCTCTGTCAATCCAGATGATATCAATGTAGAC 3634

Qy 81 GlnAspLeuValGlyTyrPProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100

Db 3635 CAAGACCTTGTGGGTGGCGCGCTCCGCAAGTAGCGGCTCAATTGACACCTGCACCTTGC 3694

Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

Db 3695 GGCTCTCTCGGACCTTTACCTGGTACAGGACACGCCGATGTCAATTCGGTGGCGCGGG 3754

Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140

Db 3755 GGTGATAGCAGGGGAGCTCTGTGCGCCCGGGCCCATTTCTTACTTGAAGGCTCTCTCG 3814

Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160

Db 3815 GGGGGTCCGCTGTGTGCCCCCGGGGACCGCGTGGGCATATTAGGGCCCGCGGTGTGC 3874

Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180

Db 3875 ACCCGTGGAGTGGCTTAAGCGGTGAGCTTTATCCTCTGTGGAGAACCTAGACAAACATG 3934

Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200

Db 3935 AGGTCCCCGGTGTTCACGGATAACTCTCTCCACCAGTATGTCGCCAGAGTTCACAGTG 3994

Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220

Db 3995 GCTCACCTCCATGCTCCACAGCGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 4054

Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

Db 4055 GCTCAGGCTATTAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGCTTTGGT 4114

Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260

Db 4115 GCTTACATGTCCAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 4174

Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280

Db 4175 ACCACTGGCAGCCCCCATCACCTACTCCACCTACCGCAAGTTCCTTGCAGCGGGGTGC 4234

Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300

Db 4235 TCGGGGGCGCTTATGACATAATAATTGTGACAGTGCCTCCACGGATGCCACATCC 4294

Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGlnThrAlaGlyAlaArgLeuThrVal 320

Db 4295 ATCTTGGGCATCGGCACCTGCTTTGACCAACGACAGACTGCGGGGGGAGACTCGTGTG 4354

Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340

Db 4355 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGAGGT 4414

Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360

Db 4415 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAAATC 4474

Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380

Db 4475 AAGGGGGGGAGACATCTCATCTTCTGTCATTAAGAAGAGTGCAGCAACTGCCGCA 4534

Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400

Db 4535 AAGCTGTGCGCATTTGGGCATCAATGCCGTGGCTTACTACCGCGGTCTTGAAGTCCGTC 4594

Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420

Db 4595 ATCCCGACAGCGGGAGTGTGTGCTGTGCAACCGATGCCCTCTATGACCGGTATACC 4654

Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

Db 4655 GCGCATCTTCGATCGGTGATAGATGCAATACGTGTGTGCACCCAGACAGTGCATTTCCAGC 4714

Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460

Db 4715 CTTGACCTTACTTCAACATTCAGATTCACGTTCGCCAGGATGCTGTCTCCGCACCT 4774

Qy 461 GlnArgAspGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480

Db 4775 CAACCTCGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGG 4834

```
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GAGCGCCCTCCGGCATGTTGCAATCGCTCGCTCTGTGAGTGCATGACGAGGCTGT 4894
Qy 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGTATGAGCTACGCCCGCGGAGACTACAGTTAGGTAGGTACGAGGACACACC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrPheTyrGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTCCCGGTGTCCAGGACCATCTTGAATTTTGGGAGGCGCTCTTTACAGGCGCTC 5014
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATATAGTCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 5075 CTGGTAGCTACCAAGCACCGTGTGGCTAGGCTCAAGCCCTCCCCCATCGTGGGAC 5134
Qy 581 GlnMetTrpLysCysIleuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCACCCCTCATGGGCCAACCCCTGCTA 5194
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGCGGCTGTTCAGATGAATCACCCTGACGACCCAGTCAACCAATACATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGTCATGTGGCGGACCTGGAGGTGCTCACGAGCACCTGGGTGCTGTGGCGG 5314
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGTGCTGTGGCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTGCTCTGTTCGGGAAGCGGCAATCATACCTCACAGGGAAGTCTCTTACCAGAGTTC 5434
Qy 681 AspGluMetGluGluCys 686
Db 5435 GATGAGATGAAGAGTGC 5452

RESULT 12
AR166930 LOCUS AR166930 9379 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6284249.
ACCESSION AR166930
VERSION AR166930.1 GI:16243325
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 9379)
AUTHORS Barban,V.
TITLE Fusion polypeptide having the C protein and E1 protein of hepatitis C virus
JOURNAL Patent: US 6284249-A 1 04-SEP-2001;
FEATURES Location/Qualifiers
source 1..9379
/mol_type="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3.45e-200 Length: 9379
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0
```

```
US-09-930-591-2 (1-686) x AR166930 (1-9379)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3395 CTGGCGCCCATCACGGCGTAGCCCCAGCAGCAAGGGGCGCTCTAGGGTGCATAATCACC 3454
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3455 AGCCTAACTGGCCGGGCAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3514
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db 3515 GCCCAAACTTCTGGCAACGTGCATCAATGGGTGTCTGGACTGTCTACACGGGGCC 3574
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3575 GGAACGAGGACCATCGCTCACCCAAAGGTCTCTGTCAATCCAGATGTATACCAATGTAGAC 3634
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3635 CAAGCACTTGTGGGTGGCGGCTCCGCAAGGTAGCGCTCATTCACACCTTCGACTTGC 3694
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 3695 GGCTCTCGGACCTTTAGCTGTGTACAGAGGACGCGCATGTCAATCCCGTGGCGGGCGG 3754
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3755 GGTGATAGCAGGGGAGGCTGTGTGCGCCGCGCCCATTTCTACTTGAAGGCTTCTCG 3814
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGGGTGCGTGTGTGTCGCCGCGGACGCGGTGGCATATTTAGGCGCGGTGTGC 3874
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCCTGTGAGTGGCTAAGCGGTGACATTATCCCTGTGGAGAACTAGAGCAACACCATG 3934
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 3935 AGTCCCGGGTGTTCACGGATACTCTCTCCACAGTAGTCCCCCAGAGCTTCAGGTG 3994
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3995 GCTCACTCCATGTCTCCACAGGACGCGGCAAAAGCAACCAAGGTCCCGGCTGCATATCA 4054
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGTCGCAACTGGGCTTGGT 4114
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCCAGGCTCATGGGATCGATCCTAAATCAGGACCGGGGTGAGAACAAATT 4174
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
Db 4175 ACCACTGGAGCGCCCATCACGTACTCCACCTACGCAAGTTCCTTCCGACGCGGGTGC 4234
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGCGCTTATGACATAATATTGTGACGAGTCCACTCCACGGATGCCACATCC 4294
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnIleAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGCACTGGGCACTGTCTTTGACCAAGCAGAGACTGGCGGGGCGAGACTGGTGTG 4354
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCCCATCCCAACATCGAGAGGT 4414
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4414
```


Db 4055 GCTCAGGGCTATAAGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4115 GCTTACATGTCGAAGGCTCATGGATCGCATCTTAAACATCAGGACCGGGTGAGACAATT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGCAGCCCATCATCAGTACTCCACTACGGCAAGTTCCTTGCCGACGGCGGTG 4234
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGGGCTTATGACATAATAATTGTGACGAGTGCATCCACGGATGCCATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGATCGGCAGTCTGCTTGACCAAGACAGACTCGGGGGGAGAGTGGTGTG 4354
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGCCACCGCCACCCCTCGGGCTCGTCACTGTGCCCATCCCAACATCGAGGAGGTT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyValAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTCTCCACCGAGAGATCCCTTTTACGGCAGAGCTATCCCTCCGAAGTAAATC 4474
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaIle 380
Db 4475 AAGGGGGGAGACATCTCTCTCTCTCAAGAGAGTGCAGCACTCGCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrArgGlyLeuAspValSerVal 400
Db 4535 AAGCTGTGCTATGGGCATCAATCGGTGGCTCTACTACGGGTCTTACGGTTCGGTC 4594
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCGACGAGCGGATGTTGCTGCTGGCAACCGATGCCCTCATGCCGTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GCGGACTTCGACTCGGTAGACTGCAATACGTGTGTCAACGAGACAGTCGATTCACG 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCTTACCTTCCATTGAGACATACAGCTCCCCAGAGTCTCTCCCGCAT 4774
QY 461 GlnArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4775 CAACGTCGGGGCAGGACTGCGAGGGGAGCCAGGCATCTACAGATTTGTCGACCGGG 4834
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GAGCGCCCTCCGGCAGTTCGACTCGTCCGTCTCTGTGAGTGTCTATCAGCAGGCTGT 4894
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGTATGAGTCAAGCCCGCGAGACTACAGTTAGGCTACAGGGTACATGACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTCCCGGTGTCAGGACCATCTTGAATTTTGGAGGGCGCTTTACAGGCCCTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
Db 5075 CTGCTAGCGTACCAAGCAGCCGTGTGGCTAGGGCTCAAGCCCTCCCTCCATCGTGGAC 5134
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGTGTGTGATTCCTCCCTCAAGCCCAACCTCTCATGGGCAACACCCCTGCTA 5194

QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGCGCTGTTCAGATGAATCACCTCGACCCAGTCACCAATAATCATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGTCATGTCCGGCAGCTCGAGTGTCTACAGAGCACCTGGTGTCTGTTGGGGC 5314
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTCGGTGCTTTGGCGCGTATTGCTGTCAACAGCTCGGTGTCATAGTGGCAGG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTCGTCTGTCCGGGAAGCCGCAATCATACCTCAGAGGAAGTCTCTACCGAGATTC 5434
QY 681 AspGluMetGluGluCys 686
Db 5435 CATGAGATGGAAGAGTGC 5452

RESULT 14
AR176483
LOCUS AR176483 9401 bp DNA linear PAT 17-DEC-2001
DEFINITION Sequence 9 from patent US 6312889.
ACCESSION AR176483
VERSION AR176483.1 GI:17918838
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 9401)
AUTHORS Houghton,M., Choo,Q.-J. and Kuo,G.
TITLE Combinations of hepatitis c virus (HCV) antigens for use in
immunassays for anti-HCV antibodies
JOURNAL Patent: US 6312889-A 06-NOV-2001;
FEATURES Location/Qualifiers
source
1..9401;
/organism="unknown"
/mol_type="unassigned DNA"

Alignment Scores:

Pred. No.: 3,46e-200 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 6 Gaps: 0

US-09-930-591-2 (1-686) x AR176483 (1-9401)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3417 CTGGCGCCCATCACGGCGTACGCCAGACAGAGGGGCTCTTAGGGTGCATATATCACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCTTAATGCGCGGACAAACCAAGTGGAGGTCAGATGTGTCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCTCTGGCAACGTCATCAATGGGTGTGTGCTGCTACACGGGCCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAACGAGACATCGCGTCAACCAAGGCTCTGTCTATCCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGGTGGCGCTCCGCAAGGTAGCGCTCATTGACACCTGCTGCTGCTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120

3717 GCTCTCTCGACCTTTACCTGGTCAAGGAGCAGCCGATGCTATCCGTCGCGCGCGG 3776
 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 3777 GGTATAGCAGGGGAGCCTCTCTCGCGCGCGCGCCATTTCTACTTGAAGGCTCTCTCG 3836
 141 GlyGlyProLeuLeuLeuProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 3837 GGGGGTCCGCTGTGTGCCCCCGCGGACGCGCGTGGGCAATATTAGGCGCGCGGTGTC 3896
 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 3897 ACCCGTGGAGTGGCTAAGGCGGTGCATTTATCCCTGTGAGAACCTTAGAGACAACCATG 3956
 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
 3957 AGGTCCCGGGTTCACGGATACTCTCTCCACCATGTAGTCCCGCAGAGCTTCCAGGTG 4016
 201 AlaHisLeuHisAlaProThrGlySerGlySerThrLysValProAlaAlaTyrAla 220
 4017 GCTCACCTCCATGCTCCACAGGAGCAGCGCAAGACCAAGTCCCGGCTGCATATGCA 4076
 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 4077 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTCTGCAACACACTGGGCTTTGGT 4136
 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaGlyThrIle 260
 4137 GCTTACATGTCAGAGGCTCATGGATCGATCCTAACATCAGGACCGGGTGAACAAT 4196
 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 4197 ACCACTGGCAGCCCATCACTACTCACCTACGCAAGTTCCTTTCGCGACGCGGGTGC 4256
 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 4257 TCGGGGGCGCTTATGACATAAATTTGTGACGAGTGCCACTCCAGGATGCCACATCC 4316
 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 4317 ATCTTGGCATCGGCACTGTCTTGACCAAGAGACTCGCGGGCGAGACTGGTTGTG 4376
 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 4377 CTCGCCACCGCCACCTCCGGGCTCGCTCACTGTGCCCTCCCAACATCGAGGAGGT 4436
 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 4437 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGCTATCCCTCCGAAGTAATC 4496
 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 4497 AAGGGGGGAGACATCTCATCTTCTGTCATCAAGAAGAGTGCACGCAACTCGCGCGCA 4556
 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 4557 AAGCTGTGCGATTTGGGATCATGCTGCGCTACTACCGCGTCTTGACGTGTCGCTC 4616
 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 4617 ATCCCGACCAAGCGCGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATAC 4676
 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 4677 GCGACCTTCGACTCGGTGATAGACTGCATATGCTGTGTACCCAGACAGCTCGATTTACG 4736
 441 LeuAspProThrThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 4737 CTTGACCTCACTTTCACCAITTAGACCAATCACGCTCCCGCAGGATGCTGTCTCCGCACT 4796
 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480

4797 CAACGTCGGGCGAGGACTGGCAGGGGAGCCAGGCATCTACAGATTGTTGGCACCGGG 4856
 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 4857 GAGCCGCCCTCCGGCATGTTGACTCGCTCTCTGAGTCTATGACGCGAGGCTGT 4916
 501 AlaTptTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 4917 GCTTGGTATAGCTCACGCCCGCGAGACTACAGTTAGCTACGAGCGTACATGAACACC 4976
 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 4977 CGGGGGCTTCGCTGTGCGAGGACCATCTTGAATTTGGGAGGGCGTCTTTACAGGGCTC 5036
 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 5037 ACTCATATAGATGCCACTTTCTATCCAGACAAGACAGAGTGGGAGAACCTTCTTAC 5096
 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
 5097 CTGCTAGCGGTACCAAGCCACCGTGTGCTAGGCTCAAGCCCTCCCGCATCGTGGAC 5156
 581 GlnMetTptLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 5157 CAGATGTGAAGTGTGTTGATTCGCTCAAGCCACCTCCAGGGCCCAACCCCTGCTA 5216
 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 5217 TACAGACTGGCGCTGTTTCAGAAATCAACCTCAGCCACCCAGTCACCAATACATC 5276
 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrThrProValLeuValGly 640
 5277 ATGACATCATGTCGGCGACCTGAGTCTCAGAGCACCTGGTGGTGGTGGCGGC 5336
 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValIleValGlyArg 660
 5337 GTCTGCTGCTGCTTGGCGCGCTATTGCTGTCAACAGGCTGCGTGGTCATAGTGGCAGG 5396
 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 5397 GTCTGCTTGTTCGGGAAGCGGCAATCATCTACAGGGAAGTCTCTACCGAGATTC 5456
 681 AspGluMetGluGluCys 686
 5457 CATGATGGAAGAGTGC 5474

RESULT 15
 E66593
 LOCUS Hepatitis C virus asialoglycoprotein. 9401 bp DNA linear PAT 18-JUN-2001
 DEFINITION
 E66593
 ACCESSION
 E66593.1 GI:13026028
 VERSION JP 1999071395-A/1.
 KEYWORDS Hepatitis C virus
 SOURCE Hepatitis C virus
 ORGANISM
 Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.
 1 (bases 1 to 9401)
 Robert,O.R., Frank,M., Kent,B.S., Barbara,A.J. and John,A.H.
 Hepatitis C virus asialoglycoprotein
 Patent: JP 1999071395-A 1 16-MAR-1999;
 CHIRON CORP
 COMMENT
 OS Hepatitis C virus
 PN JP 1999071395-A/1
 PD 16-MAR-1999
 PF 14-APR-1998 JP 1998103178
 PR 08-NOV-1990 US 611.965,09-NOV-1990 US 611.419 PR
 13-SEP-1991 US 759.880
 PI ROBERT O RARUSUTON,FRANK MARCUS,KENT B SUDIAMU, PI BARBARA A
 JABASU,JOHN A HALL
 PC
 C07K14/18,A61K39/29,G01N33/576//C12N15/09,C12P21/02,C12P21/02, PC
 C12R1:91),

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:56 ; Search time 133 Seconds
(without alignments)
2862.378 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITAYAGTGLGCIIT.....PALIPDEVLYREFDEMEEC 686

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US0930591/runat 15092004.164647 7129/app query.fasta 1.839
-DB=issued Patents NA -QRM=fastap -SUFFIX=rni -NINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNIT=Bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US0930591@cgn 1.1 59 @runat 15092004.164647 7129 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEX=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/prodata/2/ina/5A COMB.seq.*
2: /cgn2_6/prodata/2/ina/5B COMB.seq.*
3: /cgn2_6/prodata/2/ina/6A COMB.seq.*
4: /cgn2_6/prodata/2/ina/6B COMB.seq.*
5: /cgn2_6/prodata/2/ina/PCBUS COMB.seq.*
6: /cgn2_6/prodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3574	98.8	7310	3	US-08-444-818-74
2	3574	98.8	9379	3	Sequence 74, Appl
3	3574	98.8	9379	4	Sequence 1, Appl
4	3574	98.8	9401	1	US-09-916-359-1
5	3574	98.8	9401	1	US-07-910-760-9
6	3574	98.8	9401	1	US-08-440-519-9
7	3574	98.8	9401	4	US-08-440-549-9
8	3571	98.7	6785	3	US-08-823-895A-25
9	3570	98.7	8316	3	US-08-444-818-65
10	3570	98.7	9185	3	US-08-444-818-88
11	3570	98.7	9185	3	US-08-444-818-122
12	3565	98.5	2058	4	US-08-444-818-123
					Sequence 2, Appl

13	3565	98.5	2058	4	US-09-881-654-1	Sequence 1, Appl
14	3565	98.5	8987	3	US-08-444-818-137	Sequence 137, App
15	3555	99.3	5360	3	US-08-444-818-53	Sequence 53, Appl
16	3551	98.1	9646	3	US-08-811-566-1	Sequence 1, Appl
17	3551	98.1	9646	4	US-09-034-756-1	Sequence 1, Appl
18	3551	98.1	12980	3	US-08-811-566-5	Sequence 5, Appl
19	3551	98.1	12980	4	US-09-034-756-5	Sequence 5, Appl
20	3544	98.0	9599	3	US-09-014-416-2	Sequence 2, Appl
21	3544	98.0	9599	3	US-09-014-416-6	Sequence 6, Appl
22	3538	97.8	9401	2	US-08-432-693-1	Sequence 1, Appl
23	3538	97.8	9416	3	US-08-811-566-19	Sequence 19, Appl
24	3538	97.8	9416	4	US-09-034-756-9	Sequence 19, Appl
25	3535	97.7	9379	3	US-08-444-818-176	Sequence 176, App
26	3532	97.6	9401	5	PCT-US91-02225-9	Sequence 9, Appl
27	3504	96.8	9416	4	US-08-823-895A-26	Sequence 26, Appl
28	3504	96.8	9416	4	US-10-104-966-13	Sequence 13, Appl
29	3429	94.8	7989	4	US-09-539-601-10	Sequence 10, Appl
30	3429	94.8	8001	4	US-09-539-601-7	Sequence 7, Appl
31	3426	94.7	8637	4	US-09-539-601-4	Sequence 4, Appl
32	3426	94.7	8649	4	US-09-539-601-13	Sequence 13, Appl
33	3426	94.7	11076	4	US-09-539-601-1	Sequence 1, Appl
34	3423	94.6	8001	4	US-09-539-601-22	Sequence 22, Appl
35	3420	94.5	6039	1	US-08-324-977-11	Sequence 11, Appl
36	3420	94.5	6039	2	US-08-384-616-11	Sequence 11, Appl
37	3420	94.5	6039	2	US-08-904-686A-11	Sequence 11, Appl
38	3420	94.5	6039	3	US-09-315-850-11	Sequence 35, Appl
39	3420	94.5	7863	1	US-08-324-977-35	Sequence 35, Appl
40	3420	94.5	7863	2	US-08-384-616-35	Sequence 35, Appl
41	3420	94.5	7863	2	US-08-904-686A-35	Sequence 35, Appl
42	3420	94.5	7863	3	US-09-315-850-35	Sequence 35, Appl
43	3420	94.5	7917	1	US-08-324-977-31	Sequence 31, Appl
44	3420	94.5	7917	2	US-08-384-616-31	Sequence 31, Appl
45	3420	94.5	7917	2	US-08-904-686A-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1

US-08-444-818-74
; Sequence 74, Application US/08444818
; Patent No. G550087
; GENERAL INFORMATION:

APPLICANT: Chien, David Y.
APPLICANT: Rutter, William J.
TITLE OF INVENTION: NANBV Diagnostics and Vaccines
NUMBER OF SEQUENCES: 777
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: CA
COUNTRY: USA
ZIP: 94608-2916
COMPUTER READABLE FORM:
; Sequence 74, Floppy disk
MEDIUM TYPE: IBM PC compatible
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444.818
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/403.590
FILING DATE: 14-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Harbin, Alisa A.
REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 74:

3695	Db	GGCTCTCTCGGACCTTTACTCTGGTACAGAGCGACGGCGATGCTATTCCTCGTGGCCGGGG	3755
121	Qy	GLYAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
3755	Db	GGTGATAGCAGGGGAGCCTGTGTGCCCGCGCCATTTCTTACTTTGAAAGGCTCTCTCG	3814
141	Qy	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
3815	Db	GGGGGTCCGTGTGTGCCCGGGGCGACCGCGTGGCGCATTTATGGGCGCGGTGTGC	3874
161	Qy	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
3875	Db	ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCCTAGAGACAACCATG	3934
181	Qy	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
3935	Db	AGGTCCCGGTGTTCACGATAACTCTCTCCACCACTAGTGGCCCAAGAGCTTCCAGGTG	3994
201	Qy	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
3995	Db	GCTCACCTCCATGCTCCACAGGCGAGCGGCAAAAGCACCAAGGTCCCGGTGCATATGCA	4054
221	Qy	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaIleThrMetGlyPheGly	240
4055	Db	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGCTGCAACACTGGGCTTTGGT	4114
241	Qy	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
4115	Db	GCATTACATGCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATT	4174
261	Qy	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
4175	Db	ACCACTGGCAGCCCCATCAGTACTCCACCTAGGCAAGTTCCCTGCGCAGCGGGGTGC	4234
281	Qy	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
4235	Db	TCGGGGGGCGCTTATGACATAAATAATTGTGACAGTGGCACTCCACGGATGCCACATCC	4294
301	Qy	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
4295	Db	ATCTTGGGCATGGCACTGTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGTTGTG	4354
321	Qy	LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal	340
4355	Db	CTCGCCACCGCCACCCCTCGGGGTCCGTCACTGTGTGCCCATCCCAACATCCAGAGAGTT	4414
341	Qy	AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle	360
4415	Db	GCCTGTTCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTATTC	4474
361	Qy	LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla	380
4475	Db	AAGGGGGGAGACATCTCATCTTCTGTGATTTCAAAGAAGAAGTGGCAGCAACTCGCGCGCA	4534
381	Qy	LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal	400
4535	Db	AAGCTGTGCGATTGGCGCATCAATCCGTGGCCCTACTACCGCGTCTTGACGTGTCCGTC	4594
401	Qy	IleProThrSerGlyAspValValValAlaAlaThrAspAlaLeuMetThrGlyPheThr	420
4595	Db	ATCCGACCAAGCGCGATGTGTGCTGTGGCAACCGATGCCCTCATGACCGGTATATCC	4654
421	Qy	GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer	440
4655	Db	GGCGCACTTCGCACTCGGTGATAGACTGCAATACGTGTGTCCCCAGACAGTCGATTTGAGC	4714
441	Qy	LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr	460
4715	Db	CTTGACCTTACCTTCACTATTGAGACAAATCAGCGTCCCGCAGGATGTGTCTCCCGCACT	4774
461	Qy	GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly	480
4775	Db	CAACGCTCGGGGACGACTGGCAGGGGAAGCCAGGCATCTACAGATTTGTGGCACCGGGG	4834

QY 481 GluArgProSerGlyMetPheAspSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 4835 GAGCGCCCTCCGCGCATGTCGACTCGTCGCTCTGTGAGTGCTATGACGAGGCTGT 4894
 QY 501 AlaTrpTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4895 GCTTGGTATGAGCTACGCGCCGCGAGACTACAGTTAGGTACAGCGTACATGAACACC 4954
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheTrpGlyLeu 540
 Db 4955 CCGGGCTTCCGCTGTCAGACCACTTGAATTTTGGGAGGCGTCTTACAGGCTC 5014
 QY 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5074
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 5075 CTGGTAGCGTACCAAGCCAGCTGTGGCTAGGGCTCAAGCCCTCCCATCTGGGAC 5134
 QY 581 GlnMetTrpLysCysLeuLeuArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5135 CAGATGTGAAGTGTGTTGATTGCGCTCAAGGCCACCCCTCCATGGGCCCAACACCCCTGCTA 5194
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 5195 TACAGACTGGGCGCTGTTCAGATGAATACCTGACGACCCAGTCAACCAATACATC 5254
 QY 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640
 Db 5255 ATGACATGTCATGTCGGCGACCTGAGGTGTCACAGACACCTGGGTGCTCGTTGGCGC 5314
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 5315 GTCTGGTGTGTTTGGCGCGTATGCTGTCAACAGGCTCGGTGTCATAGTGGGCGAG 5374
 QY 661 IleValLeuSerGlyLysProAlaIlePheProAspArgGluValLeuTyrArgGluPhe 680
 Db 5375 GTCTGTGTCGGGAGCCGCGCAATCATACCTGACAGGAGTCTCTTACCGAGAGTTC 5434
 QY 681 AspGluMetGluGluCys 686
 Db 5435 GATGAGATGGAAGATGC 5452

RESULT 3
 US-09-916-359-1
 ; Sequence 1, Application US/09916359
 ; Patent No. 6538123
 ; GENERAL INFORMATION:
 ; APPLICANT: Veronique Barban
 ; TITLE OF INVENTION: VACCINE COMPOSITION FOR PREVENTING OR
 ; TITLE OF INVENTION: TREATING C HEPATITIS
 ; FILE REFERENCE: PWC97-03A
 ; CURRENT APPLICATION NUMBER: US/09/916.359
 ; CURRENT FILING DATE: 2001-07-26
 ; PRIOR APPLICATION NUMBER: 09/388,874
 ; PRIOR FILING DATE: 1999-09-02
 ; PRIOR APPLICATION NUMBER: 97/02,887
 ; PRIOR FILING DATE: 1997-03-06
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 9379
 ; TYPE: DNA
 ; ORGANISM: Virus
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (320)...(9352)
 US-09-916-359-1

Alignment Scores:
 Pred. No.: 0 Length: 9379

Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: Gaps: 0
 US-09-930-591-2 (1-686) x US-09-916-359-1 (1-9379)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3395 CTGGCGCCCATCACGCGCTACGCCAGCAGACAAGGGCCCTCCTAGGTGTCATAATCAC 3454
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 3455 AGCCTAACTGGCCGGGACAAAACCAAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3514
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 3515 GCCCAACCTTCTGGCAACGTCATCAATGGGTGTCTGACTGTCTACACGGGGCC 3574
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 3575 GGAACGAGGACCATCGCGTCAACCAAGGCTCTGTCTATCCAGATGTATACCAATGTAGAC 3634
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3635 CHAGACCTTGTGGGTGGCCGCTCGCAAGTAGCGCTCATTTGACACCTCGACTTGC 3694
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 3695 GCGTCTCGGACCTTTTACCTGTCACGAGGACCCGATGTCAATTCGCTGCGCGCGCGG 3754
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 3755 GGTGATAGCAGGGGCGCTGCTGTCGCCCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG 3814
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 3815 GGGGTGCTGCTGTGTGTCGCGCGGCGCACGCCGTGGGCATATTTAGGGCGCGGTGTGC 3874
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 3875 ACCCTGGAGTGGCTAAGCGGTGACCTTTATCCTGTGGAGAACCTTAGACACCAACATG 3934
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 3935 AGGTCCCCGGTGTTCACGGATACTCTCTCCACCAGTAGTGTGCCCGCAGAGCTTCCAGGTG 3994
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 3995 GCTCACCTCCATGCTCCACAGGCGCGCAAGACCAAGAGTCCCGGCTGCATATGCA 4054
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 4055 GCTCAGGCTATAGTGTCTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT 4114
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4115 GCTTACATGTCCAAGGCTCATGGGATCATCTTAACATCAGGACCGGGGTGAGAACAT 4174
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 4175 ACCACTGGCAGCCCATCCGCTACTCCACCTACGGCAAGTTCCTTTCGCGAGGGGGTGC 4234
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4235 TCGGGGGCGCTTATGACATAATAATTGTGACAGTGCCACTCCACGGATGCCACATCC 4294
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 4295 ATCTTTGGGCATCGGCACCTGCTCTTGACCAACAGACAGACTGGGGGGGCGAGACTGGTGTG 4354
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340

5435 GATGAGATGGAGAGTGC 5452

Db

RESULT 4

US-07-910-760-9

Sequence 9, Application US/07910760

Patent No. 5683864

GENERAL INFORMATION:

APPLICANT: Houghton, Michael

APPLICANT: Choo, Qui-Lim

APPLICANT: Kuo, George

TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)

TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation

STREET: P.O. Box 8097 (Int. Prop. R-440)

CITY: Emeryville

STATE: CA

COUNTRY: U.S.A.

ZIP: 94662-8097

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/910,760

FILING DATE: 07-JUL-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Blackburn Esq., Robert P.

REGISTRATION NUMBER: 30,447

REFERENCE/DOCKET NUMBER: 0101.002

TELECOMMUNICATION INFORMATION:

TELEPHONE: (510) 601-2702

TELEFAX: (510) 655-3542

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 9401 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 342..9374

FEATURE:

NAME/KEY: misc_feature

LOCATION: 366

OTHER INFORMATION: /note= "This amino acid position

OTHER INFORMATION: can also be Arg."

FEATURE:

NAME/KEY: misc_feature

LOCATION: 372

OTHER INFORMATION: /note= "This amino acid position

OTHER INFORMATION: can also be Thr."

FEATURE:

NAME/KEY: misc_feature

LOCATION: 867

OTHER INFORMATION: /note= "This amino acid position

OTHER INFORMATION: can also be Thr."

FEATURE:

NAME/KEY: misc_feature

LOCATION: 1341

OTHER INFORMATION: /note= "This amino acid position

OTHER INFORMATION: can also be Val."

FEATURE:

NAME/KEY: misc_feature

LOCATION: 2148

OTHER INFORMATION: /note= "This amino acid position

OTHER INFORMATION: can also be Ile."

FEATURE:

NAME/KEY: misc_feature


```

/ LOCATION: 2883
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Asn."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3681
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 3690
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Thr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4167
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Leu."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4323
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4701
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Tyr."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 4752
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 5970
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6183
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be His."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6186
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Cys."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 6402
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Val."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7386
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ser."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7494
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7497
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Ala."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 7845
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Phe."

```

```

/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 8409
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9102
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Gly."
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 9327
/ OTHER INFORMATION: /note= "This amino acid position
/ OTHER INFORMATION: can also be Pro."
/ US-07-910-760-9
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
US-09-930-591-2 (1-686) x US-07-910-760-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3417 CTGGCGCCCATCACGCGCTAGCCGACAGACAAGGGCCCTCCTAGGGTGCAATACACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 3477 AGCCTAACTGGCGGACAAAACCAAGTGAGGGTGAGGTCAGATTGTCTCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 3537 GCCCAACCTTCTGGCAACGTGCATCAATGGGTGTGTGGACTGTCTACACGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 3597 GGACGAGGACCATCGCGTCACCAAGGGTCTCTCATCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTTGTGGCTGGCCGCTCCGCAAGTAGCGCTCATTGACACCCCTGCATTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 3717 GGCTCTCGGACCTTTTACCTGGTCACGAGGACCGCATGTCTCCCGTGGCGCGCGG 3776
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 3777 GGTGATAGCAGGGGAGCGCTGCTGTGCCCCGGCCCATTTCTACTTGAAGAGGTCTCTCG 3836
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 3837 GGGGCTCCGCTGTTGTGCCCCGGGGCACGCCGTGGGCATATTTAGGGCGCGCGTGTGC 3896
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 3897 ACCGTGGAGTGGCTTAAGCGGTGACTTTATCTCTGTGGAGAACCTTAGACACACCATG 3956
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 3957 AGGTCCCGGTGTTTACGAGATAACTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 4016
QY 201 AlaHisLeuHisAlaProThrGlySerGlySerThrLysValProAlaAlaTyrAla 220
DB 4017 GCTCACCTCCATGCTCCACAGGCGGCAAAAGCACCAGGTCCCGGTGCAATATGCA 4076
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

```

Db 4077 GCTCAGGGCTAAGTGTCTAGTACTCAACCCCTCTGTGTCTGCTGCAACACTGGGCTTTGGT 4136
 Qy 241 AlaTyrMetSerLysAlaHisGlyLeuAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4137 GCTTACATGTCGAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAATT 4196
 Qy 261 ThrThrCysProIleThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 Db 4197 ACCACTGGCAGGAGCCCATCAGTACTCCACTACGGCAAGTCTCTTGCCGACCGCGGGTGC 4256
 Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4257 TCGGGGGGGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGAGATGCCACATCC 4316
 Qy 301 IleGluGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgThrVal 320
 Db 4317 ATCTTGGGATCGGCATCTGCTTGAACAGAGAGACTGCGGGGGCGAGACTGGTTGTG 4376
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 4377 CTGCCACCGCCACCCCTCGGGCTCGGTCACTGTGCCCCCATCCCAACATCCGAGGAGTT 4436
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyIleAlaIleProLeuGluAlaIle 360
 Db 4437 GCTCTGTCCACCGAGAGATCCCTTTTACGGCAAGCTATCCCTCGAAAGTAATC 4496
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 4497 AAGGGGGGAGACATCTCATCTCTGTCAATCAAGAGAGTGCAGCACTCGCGCA 4556
 Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 4557 AAGCTGTGTCATTGGGCATCAATGCGGTGCGCTACTACCGCGCTGTGACGTGTCGTC 4616
 Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 4617 ATCCCGACCGCGGATGTGCTGCTGCGCAACCGATGCCCTCATGACCGGCTATACC 4676
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 4677 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTTTCCAGC 4736
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4737 CTTGACCCCTACCTTCACTTACGATGAGCAATACGCTCCCGGAGATGCTGTCTCCCGCACT 4796
 Qy 461 GluArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 4797 CAAGCTCGGGGAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCAGCGGG 4856
 Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 4857 GAGCGCCCTCCGGCATGTTTCGACTCGCTGCTCTGTGAGTGTATGACGCGAGGCTGT 4916
 Qy 501 AlaTyrThrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4917 GCTTGGTATGAGCTACGCGCGCGGAGACATACAGTATAGGCTACGAGGTACATGAACACC 4976
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
 Db 4977 CCGGGGCTTCCGCTGCCAGGACCATTTGAATTTTGGGAGGCGCTTTTACAGGCGCTC 5036
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 5037 ACTCATATAGATGCCCATTTCTATCCAGAACAGAGAGTGGGGGAGAACCTTCCCTTAC 5096
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTyrAsp 580
 Db 5097 CTGTAGCTTACCAAGCACCGCTGTGCGTAGGGCTCAAGCCCTTCCCTCATCGTGGGAC 5156
 Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5157 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCCACTCCATCCATGGGCCCAACACCCCTGCTA 5216

RESULT 5

US-08-440-519-9
 ; Sequence 9, Application US/08440519
 ; Patent No. 5712087
 ; GENERAL INFORMATION:
 ; APPLICANT: Houghton, Michael
 ; APPLICANT: Chco, Qui-Lim
 ; APPLICANT: Kuo, George
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
 ; TITLE OF INVENTION: Antigens for Use in Immunoassays for Anti-HCV Antibodies
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: P.O. Box 8097 (Int. Prop. R-440)
 ; CITY: Emeryville
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94662-8097
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/440,519
 ; FILING DATE: 12-MAY-1995
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/910,760
 ; FILING DATE: 07-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Blackburn Esq., Robert P.
 ; REGISTRATION NUMBER: 30,447
 ; REFERENCE/DOCKET NUMBER: 0101.002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (510) 601-2702
 ; TELEFAX: (510) 655-3542
 ; INFORMATION FOR SEQ ID NO: 9:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 9401 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 342..9374
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 366
 ; OTHER INFORMATION: /note= "this amino acid position
 ; OTHER INFORMATION: can also be Arg."

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402

```

```

FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also by Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-519-9
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0
US-09-930-591-2 (1-686) x US-08-440-519-9 (1-9401)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 3417 CTGGGGCCCATCAGCGGCTACGCCCCAGACACAGGGGCTCTTAGGTCGATAATCACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThra 40
DB 3477 AGCCTAACTGGCCGGGACAAAAACCAAGTGGAGGGTGAGGTCAGATTGTCAACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 3537 GCCAAACCTTCCTGGCAACGTGTCATATGGGGTGTGGACTGTCTACACGGGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValasp 80
DB 3597 GGAACGAGGACCATCGCGTCACCCAGGGTCCTGTATCCAGATGATATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTTPProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 3657 CAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCGCTCATTCACACCTGCATTCG 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaaspValIleProValArgArg 120
DB 3717 GGCTCTCTCGGACCTTTACCTGGTCACGAGGCACGCCGATGCTCATTCCCGTGGCGCGG 3776

```

QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLeuGlySerSer 140
 Db 3777 GGTGATAGAGGGGCGACCTGCTGTGCGCCCGCCCATTTTCCTACTTGAAGGCTCCTCG 3836
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 3837 GGGGGTCCGCTGTGTGCGCGCGGGGACGCGCTGGGCATATTTAGGGCCGCGGTGTC 3896
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 3897 ACCCGTGGAGTGGCTAAGCGCGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCATG 3956
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 3957 AGGTCCCGGTGTTCACGGATTAATCTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 4016
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLeuSerThrLysValProAlaAlaTyrAla 220
 Db 4017 GCTCACCTCCATGCTCCACAGCGAGCGGCAAAAGCACCAAGGTCGCGGCTGCATATGCA 4076
 QY 221 AlaGlyGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 4077 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 4136
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4137 GCTTACATGTCCAAGGCTCATGGATCGATCCCTAACATCAGGACCGGGTGAGAACAAAT 4196
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaPheGlyLysCys 280
 Db 4197 ACCACTGGACGCCCATCAGTACTCCACCTACGGCAAGTTCCTTCCGACGCGGGTGC 4256
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4257 TCGGGGGGCGCTTATGACATATAATTTGTGAGAGTGGCCATCTCCACGGATGCCACATCC 4316
 QY 301 IleuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 4317 ATCTTGGGCATCGGCATGTCCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTTGTG 4376
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 4377 CTGCGCACCGCCACCCCTCGGGCTCGCTCAGTGTGCCCATCCCAACATCGAGGAGGT 4436
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 4437 GCTCTGTCACCCAGCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 4496
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGGAAGTGGCGAGCACTCGCCGCA 4556
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 4557 AAGTGTGTCATTTGGGCATCATCGCTGGCTACTACCGCGTCTTGACGTGTCCGTC 4616
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 4617 ATCCCGACCGACGCGCATGTTGCTGCTGGCAACCGATGCCCTCATGACCGGCTATACC 4676
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 4677 GCGACCTTCGACTCGGTGATAGATCGCAATACGTGTCTCCACAGACGTCGATTTCCAGC 4736
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4737 CTTGACCTCACTTCCACATTGACACAATCAGCTCCCGGATGCTGTCTCCGCACT 4796
 QY 461 GlnArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly 480
 Db 4797 CAACGTCGGGCGAGACTGCGAGGGGAAGCCAGGCATCTACAGATTGTGGACCGGGG 4856
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500

Db 4857 GAGCCCCCTCCGGCATGTTGCACTGCTGCTCTCTGTGAGTGTCTATGACGAGGCTGT 4916
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuAlaTyrMetAsnThr 520
 Db 4917 GCTTGGTATGAGTCAACGCCCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4976
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
 Db 4977 CCGGGGCTTCCCGTGTGCGAGACCATCTTGAATTTTGGAGGGGCTTTTACAGGCTC 5036
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 5037 ACTCATATAGATGCCCACTTTCTATCCAGACAAGCAGAGTGGGAGAACCTTCTTAC 5096
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
 Db 5097 CTGGTAGGGTACCAAGCCACCTGTGCGCTAGGCTCAAGCCCTCCCTCCCATCGTGGGAC 5156
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5157 CAGATGTGAAGTGTTCATTCGCTCAAGCCCTCCATGGGCCAACACCCCTGCTA 5216
 QY 601 TyrArgLeuGlyAlaValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 5217 TACAGACTGGGCGCTGTTCAGAAATGAAATCAACCTGACGACCCAGTCAACCAATACATC 5276
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 5277 ATGACATCATGTGCGCGACCTGGAGTGTGTCAGACACCTGGGTGCTCTTGCGCGC 5336
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 5337 GTCTGGTGTCTTTGGCGCGGTATTGCTGTCAACAGAGCTGCTGTGTCATAGTGGCAGG 5396
 QY 661 IleValLeuSerGlyProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 5397 GTCTGTCTTGTCCGGAGCGCAATCATCTGACAGGGAAGTCTCTTACCGAGATTC 5456
 QY 681 AspGluMetGluGluCys 686
 Db 5457 GATGAGATGGAAGATGC 5474

RESULT 6
 US-08-440-549-9
 ; Sequence 9, Application US/08440549
 ; Patent No. 6312889
 ; GENERAL INFORMATION:
 ; APPLICANT: Houghton, Michael
 ; APPLICANT: Choo, Qui-Lim
 ; TITLE OF INVENTION: Kuo, George
 ; TITLE OF INVENTION: Combinations of Hepatitis C virus (HCV)
 ; NUMBER OF SEQUENCES: 12
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Chiron Corporation
 ; STREET: P.O. Box 8097 (Int. Prop. R-440)
 ; CITY: Emeryville
 ; STATE: CA
 ; COUNTRY: U.S.A.
 ; ZIP: 94662-8097
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/440,549
 ; FILING DATE: 12-MAY-1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 07/910,760
 ; FILING DATE: 07-JUL-1992

ATTORNEY/AGENT INFORMATION:
NAME: Blackburn Esq., Robert P.
REGISTRATION NUMBER: 30,447
REFERENCE/DOCKET NUMBER: G101.002
TELEPHONE: (510) 601-2702
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 9401 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 342..9374
FEATURE:
NAME/KEY: misc_feature
LOCATION: 366
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Arg."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 372
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 867
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1341
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2148
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ile."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 2883
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Asn."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3681
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 3690
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Thr."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4167
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Leu."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4323
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 4701
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Tyr."

NAME/KEY: misc_feature
LOCATION: 4752
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 5970
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6183
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be His."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6186
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Cys."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 6402
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Val."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7386
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ser."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7494
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7497
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Ala."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 7845
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Phe."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 8409
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9102
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Gly."
FEATURE:
NAME/KEY: misc_feature
LOCATION: 9327
OTHER INFORMATION: /note= "This amino acid position
OTHER INFORMATION: can also be Pro."
US-08-440-549-9
Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-08-440-549-9 (1-9401)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
:::|||||

Db 3417 CTGGGGCCATCAGCGGCTACGCCAGACAGCAAGGGCCCTCCTAGGGTGCATAATCAACC 3476
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
Db 3477 AGCCTAACTGGCGGGCAAAACCAAGTAGTGAGGGTGAGTCCAGATTTGTCAACTGCT 3536
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAAACTTCTGGCAACGTGCATCAATGGGGTGTCTGGACTGCTACACAGGGGCC 3596
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGAAACGAGGACCAATCGCGGTCAACCAAGGGTCCGTGTCTACAGATGTATACCAATGTAGAC 3656
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGCTGGCGGCTCCGCAAGGTAGCGGCTCATTGACACCCCTGCACCTGC 3716
Qy 101 GlySerSerAspLeuThrValThrArgHisAlaAspValIleProValArgArgArg 120
Db 3717 GGCTCTCTGGACCTTTACCTGGTCACGAGGCAAGCCGATGTCAATTCGCGTGGCGCGCG 3776
Qy 121 GlyAspGlyArgGlySerLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3777 GGTGATAGCGGGACGCTGCTGTGCCCGCGGCCAATTTCTTACTTGAAGGCTCTCTCG 3836
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3837 GGGGGTCCGTGTGTGCCCGCGGGCACCGCGTGGGCATATTTAGGGCGCGGTGTGC 3896
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3897 ACCCGTGGAGTGGTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAAACATG 3956
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3957 AGGTCCCGGTGTTCAGGATAACTCTCTCCACCACTAGTGCCCGACAGCTTCCAGGTG 4016
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 4017 GCTCACCTCCATGCTCCAGGAGCGGCAAGAACCAAGCTCCCGGTGTGATATGCA 4076
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4077 GCTCAGGGCTATAGGTGTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT 4136
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 4137 GCTTACATGTCMAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAAT 4196
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4197 ACCACTGGCAGCCCATCAGCTACTCCACCTACGGCAAGTTCCTTGGCCGACGGCGGGTGC 4256
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4257 TCGGGGGCGCTTATGACATAATAATTGTGACGAGTGCCACTCCACGGATGCCACATCC 4316
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4317 ATCTTGGGCATCGGCACCTGCTTGTGACCAAGCAGAGACTCGGGGGCGAGACTGGTGTG 4376
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4377 CTCGCCACCGCACCCCTCCGGGCTCGTCACTGTGCCCATCCCAACATCGAGGAGGT 4436
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4437 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 4496
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGGAAGTGTGGCAGCAACTCGCGCA 4556

RESULT 7

US-08-823-895A-25
; Sequence 25, Application US/08823895A
; Patent No. 6433159
; GENERAL INFORMATION:
; APPLICANT: Kevin P. Anderson
; TITLE OF INVENTION: Compositions And Methods For
; TREATMENT OF Hepatitis C Virus-Associated Diseases

```

; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESS: Jane Massey Licata, Esq.
; STREET: 66 E. Main Street
; CITY: Marlton
; STATE: NJ
; COUNTRY: USA
; ZIP: 08053
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/823.895A
; FILING DATE: March 17, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/453,085
; FILING DATE: May 30, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/945,289
; FILING DATE: September 10, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: ISPH-0203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 810-1454
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9401
; TYPE: Nucleic
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: NO
; US-08-823-895A-25

Alignment Scores:
Pred. No.: 0 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservatve: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-08-823-895A-25 (1-9401)
QY 1 MetAlaProHleThrAlaTyrAlaGlnGlnThrArgGlyLeuGlyCysIlelleThr 20
Db 3417 CTGGCGCCCATCAGCGGTACGCCAGACAGAGGGGCTCTCTAGGGTGCATAATCACC 3476
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 3477 AGCCTAATCGCGGACAAACCAAGTGGAGGTGAGTCCAGATTGTGCTCACTGCT 3536
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 3537 GCCCAACCTTCCTGGCAACGTGATCAATGGGTGTGCTGGAGTGTCTACCAACGGGCC 3596
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3597 GGACGAGGACATCGGTACCCAAAGGTCTCTCATCCAGATGTATACCAATGTAGAC 3656
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3657 CAAGACCTTGTGGGTGGCGGCTCCGACAGGTAGCGGCTCAATGACACCTGCATCTGC 3716
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 3717 GGCTCTCGGACCTTACCTGTGTACGAGGACCCGATGTCATTCCTCCGTGGCCGGCGG 3776

```


Db 4857 GAGCGCCCTCCGCAATGTTCCGACTGCTCGCTCTCTGTGAGTGCATATGACGCGAGCTGT 4916
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4917 GCTTGGTATGAGTCACGCGCGCGAGACTACAGTATAGGCTACGAGGTACATGAACACC 4976
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4977 CCGGGGCTCCCGTGTCCAGGACCATCTTGAATTTGGAGGGCGCTTTTACAGGCCCTC 5036
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5037 ACTCATATAGATGCCACTTTCTATCCAGAACAGAGAGTGGGAGAACCTTCTCTTAC 5096
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5097 CTGTAGCGTACCAAGCACCGTGTGCGCTAGGCTCAGGCCCTCCCCCATCGTGGAC 5156
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5157 CAGATGTGGAAGTGTGTTGATTTGGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 5216
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5217 TACAGCTGGCGCTGTTCAGATGAATCACCCTGACGCCACCCAGTCACCAATACATC 5276
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5277 ATGACATGCATGTGCGCGACCTGAGGTGCTCACGAGCACCTGGGTGCTCGTTGGCGGC 5336
Qy 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5337 GTCTGCTGCTTTGGCGCGGTATTGCTCTCAACAGGCTGCGTGTGTCATAGTGGCAGG 5396
Qy 661 IleValLeuSerGlyLysProAlaIlePheProAspArgGluValLeuTyrArgGluPhe 680
Db 5397 GTCTGCTTTGCGGAGCGCGCAATCATCTGACAGGAAATGCTCTCTACCGAGAGTTC 5456
Qy 681 AspGluMetGluGluCys 686
Db 5457 GATGAGATGGAAGAGTGC 5474

RESULT 8

US-08-444-818-65
; Sequence 65, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisha A.

REGISTRATION NUMBER: 33,895
REFERENCE/DOCKET NUMBER: 0110.002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (508)359-3876
TELEFAX: (508)359-3885
INFORMATION FOR SEQ ID NO: 65:
SEQUENCE CHARACTERISTICS:
LENGTH: 6785 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 3..6785
US-08-444-818-65
Alignment Scores:
Pred. No.: 0 Length: 6785
Score: 3571.00 Matches: 671
Percent Similarity: 99.42% Conservativeness: 11
Best Local Similarity: 97.81% Mismatches: 4
Query Match: 98.70% Indels: 0
DB: 3 Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-65 (1-6785)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1203 CTGGCGCCCATCAGCGCTAGCCCGCAGCAGCAAGGGGCTCTTAGGTGCATATCACC 1262
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1263 AGCCTTAACCTGGCGCGGCAAAAACCAAGTGGAGGCTGAGGTCCAGATTGTGTCAACTGCT 1322
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db 1323 GCCCAAACTTCTCTGGCAACGTGCATCATATGGGTGTGTCTGGACTGTCTACACCGGGGCC 1382
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1383 GGAAGCAGGACCATCGCGTCACCAAGGCTCCTGTCTATCCAGATGTATACCAATATAGAC 1442
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1443 CAAGACCTTGTGGGTGTGGCCGCTCCGAGAGTAGCCGCTCATTCACACCTTGCACTTGC 1502
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1503 GGCTCTCTGGACCTTATCCTGGTACAGGACGCGCGCATGTCTATCCGTCGCGCGGGGG 1562
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProLysSerTyrLeuLysGlySerSer 140
Db 1563 GGTGATAGCAGGGGAGGCTGTCTGCGCCCGGCCCATTTCTCTTGAAGGGCTCTCTCG 1622
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 1623 GGGGTCTCGCTGTGTGCGCGCGGCGCACGCGGTGGGCATATTTAGGGCCCGCGGTGTC 1682
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 1683 ACCCGTGGAGTGGCTTAAGCGGTGGACTTATCTCCCTGTGGAGAACCTAGAGAACACCATG 1742
Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 1743 AGGTCCCGGTGTTCACGGATACCTCTCTCCACAGTAGTGGCCCGCAGAGCTTCAGGGTG 1802
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 1803 GCTCACCTCCATGCTCCACAGGAGCGCAAAAGCAACCAAGGTCCCGGCTGCATATGCA 1862
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240

Db 1863 GCTCAGGGCTATAGGTCCTAGTACTCAACCCCTCTGTTGCTGCAACACACTGGGCTTTGGT 1922
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1923 GCTTACATGTCACAGGCTCATGGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAAATT 1982
Qy 261 ThrThrGlySerProIleThrTyrSerThrThrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 1983 ACCACTGGCAGCCCATCACGTAATCCACCTACGCGCAAGTTCTTTGCGGACGGGGGTGC 2042
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2043 TCGGGGGCGCTTATGACATAATAATTTGTGACGACTGCCACTCCACGGATGCCACATCC 2102
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2103 ATCTTGGGCATCGGCATGCTCTTGACCAAGCAGAGACTCGGGGGCGGACTGGTTGTG 2162
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2163 CTCGCCACCGCCACCCCTCCGGCTCCGTCACTGTGCCCATCCCAACATCGAGGAGTT 2222
Qy 341 AlaLeuSerThrThrGlyGluIleProPheThrGlyLysAlaIleProLeuGluAlaIle 360
Db 2223 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 2282
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
Db 2283 AAGGGGGGAGACATCTCATCTTGTCAATCAAGAGAGTGCAGACGACTCGCGCA 2342
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2343 AAGCTGTGCGATTCGGGCATCAATCGCGTGGCTACTACCGCGCTCTTGACGTGTCGTC 2402
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2403 ATCCGACACGGCGGATGTTGTCTGTGGCAACCGATGCCCTCATGACCGGCTATACC 2462
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2463 GCGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCCACCCAGACAGTCGATTTGAGC 2522
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2523 CTTGACCCCTACCTTCAACATTGACCAATCACGCTCCCGCAGGATGCTGCTCCCGCACT 2582
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2583 CAACGTGGGGCAGACTGGCGAGGGGAGCCAGGCATCTACAGATTGTGGCACCGGG 2642
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2643 GAGCGCCCTCCGGCATGTTCCAGTCTCGTCTGCTGTGAGTGCTATGACGAGGCTGT 2702
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2703 GCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGGCTACATGACACC 2762
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 2763 CCGGGGCTTCCCGTGTGCGAGGACCATTTGAATTTTGGGAGGGCGCTTTTACAGGCTC 2822
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 2823 ACTCATATAGATGCCACTTTCTATCCAGCAAGCAGAGTGGGAGAACCTTCCTTAC 2882
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 2883 CTGGTAGCGTACCAAGCCACCGTGTGGCTAGGGCTCAAGCCCTCCCGCATCGTGGAC 2942
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2943 CAGATGTGGAAGTGTGTTGATTCGGCTCAAGGCCCAACCTTCCATGGGCGCAACCCCTGCTA 3002

Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3003 TACAGACTGGGGCTGTTCAGAAATACACCTTGACGCGCCAGTCACCAATATACATC 3062
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3063 ATGACATGCATGTCCGCCGACCTGGAGGTCGTCAGGACACCTGGTGCTCGTTGGCGGC 3122
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3123 GTCTCTGGCTGCTTTGGCGCGTATTCCTGCTCAACAGGCTCGTGGTCATAGTGGCGAG 3182
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3183 GTCTGCTGTTCGGGAGCCGCGCATCATACTGACAGGGAAGTCCCTCTACCGAGAGTTC 3242
Qy 681 AspGluMetGluGluCys 686
Db 3243 GATGAGATGGAAGAGTGC 3260

RESULT 9

US-08-444-818-88
; Sequence 88, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508) 359-3876
; TELEFAX: (508) 359-3885
; INFORMATION FOR SEQ ID NO: 88:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8316 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8316
; US-08-444-818-88

Alignment Scores:
Pred. No.: 0 Length: 8316
Score: 3570.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5

Query Match: 98.67% Indels: 0
 DB: 3 Gaps: 0
 US-09-930-591-2 (1-666) x US-08-444-818-88 (1-8316)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 2734 CTGCGCCCATCACGGCGTACGCCAGCAGACAAGGGGCTCTAGGGTGCATAATCACC 2793
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla 40
 DB 2794 AGCTTAACCTGCGCGGACCAAAACCAAGTGGAGGTTGAGGTCCAGATTGTGTCACTGCT 2853
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 DB 2854 GCCCAAACTTCTCGCAACGTCATCAATGGGTGTGTGGACTGTCTACCAACGGGGCC 2913
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 2914 GGACGAGGACCATCGCGTACCACCAAGGTCCTGTCTCATCAGATGATACCAATGTAGAC 2973
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 2974 CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCGGCTCATTGACACCTGTGCACTTCG 3033
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 DB 3034 GGCTCTCGGACCTTACCTGTGTGTCAGAGCAGCGCGATGTCAATCCCGTGGCGCGCGG 3093
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 3094 GGTGATAGAGGGGCGGCTGTGCTGCCCGGCCCATTTCTTACTTTGAAGGCTCTCTCG 3153
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaValCys 160
 DB 3154 GGGGGTCCGCTGTGTGTCGCCCGGGGCGACCGCTGGGCATATTTAGGCGCGGCTGTGC 3213
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 3214 ACCGTGGAGTGGCTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTAGACNACCATG 3273
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 3274 AGTCCCCGGTGTTCAGGATAACTCTCTCCACCATAGTAGTCCCGCAGAGCTTCCAGGTG 3333
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 3334 GCTCACCTCATGCTCCACAGCAGCGGCAAGAACACCAAGTCCCGGTGCTCATATGCA 3393
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 3394 GCTCAGGCTATAAGGTGTGTACTCAACCCCTCTGTTCGTAACACTGGGCTTTGGT 3453
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 3454 GCATTACATGTCCAAGGCTATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAAT 3513
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 3514 ACCACTGGACCCCATCAGTACTCCACTACGGGAGTTCTTCCCGACGGCGGTGC 3573
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 3574 TCGGGGGCGGTTATGACATAATAATTTGTGACAGTGCACCTCCACGGATGCCATCC 3633
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 3634 ATCTTGGGCTATCGGCACTGTCTTGACACAGAGACTCGGGGGGAGACTGGTTGTG 3693
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 3694 CTCGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCCGAGGAGTT 3753

QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 DB 3754 GCTCTGTCCACCCAGGAGATCCCTTTTACGGCAAGCTATCCCCCTCGAAGTAAATC 3813
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 DB 3814 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAAGAAGTGGCAGCAACTCGCCGCA 3873
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 3874 AAGCTGTGCGATTTGGGCATCAATGCCGTGGCTTACTACCGCGCTTTCAGTGTCCGTC 3933
 QY 401 IleProThrSerGlyAspValValValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 3934 ATCCCGACACGGCGGATTTGTCTGTGGCAACCGATGCCCTCATGACGGGTATACC 3993
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 3994 GGCACCTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCCAGACAGTGTGATTTCAGC 4053
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 4054 CTTGACCCCTACTCTTACCATTTGAGACAATACGCTCCCGCAGGATGCTGTCTCCCGACT 4113
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 DB 4114 CRAAGTGGGGCAGGACTGGCAGGGGAGAGCCAGGCATCTACAGATTGTGGCACCGGGG 4173
 QY 481 GluArgProSerGlyMetPheAspSerValLeuGlyCysGluCysTyrAspAlaGlyCys 500
 DB 4174 GAGCGCCCTCCCGCATGTTCCGACTGCTGCTGTGTGAGTGTGTATGACGAGGCTGT 4233
 QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 4234 GCTTGGTATGAGTCCACGCCCGCGAGACTACAGTTAGGTACGAGCGTACATGACACCC 4293
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 DB 4294 CCGGGGCTTCCGTTGTCAGGACCATCTTGAATTTTGGGAGGGGCTTTTACAGGCCTC 4353
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 DB 4354 ACTCATATAGATGCCCATCTTCTATCCAGACAAGAGAGTGGGAGAGACCTTCCTTAC 4413
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
 DB 4414 CTGGTAGCTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCGCATCGTGGAC 4473
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 4474 CAGATGTGGAAGTGTGTTGATTGCGCTCAAGCCCACTCCATGGGCGCAACCCCTGCTA 4533
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 DB 4534 TACAGACTGGGCGCTGTTTCAGATCAATACCTGAGCGCACCCAGTACCACAAATACATC 4593
 QY 621 MetThrCysMetSerAlaAspLeuValValThrSerThrTrpValLeuValGlyGly 640
 DB 4594 ATGACATGCAATGTCCGCGGACCTGAGGTCGTACAGACACCTGGGTCTCGTGGCGGC 4653
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 4654 GTCTGGCTGTCTTGGCGCGGTATTCCTGTCAACAGAGCTGGGTGTCTAGTGGGCGAG 4713
 QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 4714 GTGCTTGTTCGGGAAGCCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 4773
 QY 681 AspGluMetGluGluCys 686
 DB 4774 GATGAGATGGAAGATGC 4791

RESULT 10


```

Db 4775 CAACGTGGGGAGGAGTGGCAGGGGAGCCAGGCGATCTACAGATTGTGGCAGCGGG 4834
Qy 481 GUAArgProSerGlyMetPheAspSerValLeuCyseGluCysTyrispAlaGlyCys 500
Db 4835 GAGCGGCGCCCTCCGGCATGTTCGACTCGTCGCTCTGTGAGTGTATACAGCGCTGT 4894
Qy 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAenThr 520
Db 4895 GCTTGGTATGAGCTACCGCCCGCGAGACTACAGTTAGGCTACGAGCGGTACATGAACCC 4954
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrPheGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGGCTCCCGTGTGCCAGGACCATCTTGAATTTTGGAGGGCGCTTTTACAGGCGTC 5014
Qy 541 ThrHisLeuAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGAGAGTGGGGAGAACCTTCTTAC 5074
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5075 CTGGTAGCGTACCAAGCCAGCGTGTGGCTTAGGGCTCAAGCCCGCTCCCGCATCGTGGGAC 5134
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAGGTGTGTGATCGCCTCAAGCCCGCCATCCATGGGCCAACACCCCTGCTA 5194
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGCGCTGTTCAGATGAATCAACCTGACGCCACCGCTACCAAAATACATC 5254
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 5255 ATGACATGTCATGTCCGCCGACCTGGAGGTGTCACGAGCACCTGGGTGCTCGTGGCGC 5314
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGGCTGCTTGGCCGGTATTCCTGTCAACAGGCTGCTGTCATAGTGGCGAG 5374
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTGCTGCTGTCCGGGAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 5434
Qy 681 AspGluMetGluGluCys 686
Db 5435 GATGAGATGGAGAGTGC 5452

```

```

RESULT 11
US-08-444-818-123/c
; Sequence 123, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 123:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9185 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ANTI-SENSE: YES
; US-08-444-818-123

Alignment Scores:
Pred. No.: 0 Length: 9185
Score: 3570.00 Matches: 671
Percent Similarity: 99.27% Conservative: 10
Best Local Similarity: 97.81% Mismatches: 5
Query Match: 98.67% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-123 (1-9185)
Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 5791 CTGGCGCCCATCACGGCTACGCCAGCAGCAGCAGAGGGGCTCTCTAGGGTGCATATAC 5732
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 5731 AGCCTAACTGGCGCGGACAAAACCAAGTGGAGGGTCCAGATTGTCTCAACTGCT 5672
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 5671 GCCCAAACTTCTGGCAACGTGCATCAATGGGTGTCTGACCTGTCTACACGGGCCC 5612
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 5611 GGAACGAGGACCATCGCGTCACCAAGGGTCTCTGTCTCCAGATGTATACCAATGTAGAC 5552
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 5551 CAAGACCTTGTGGGCTGGCGCTCCCGCAAGGTAGCCGCTCATTGACACCCCTGCACITGC 5492
Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db 5491 GGTCTCTCGGACCTTACCTGGTCAGAGGACGCGCGGATGTCATTCGCTGCGCGCGCGG 5432
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 5431 GGTGATAGCAGGGCAGCGCTGCTGTGCGCCCGGCGCCATTTCTCTACTTGAAGAGGTCTCT 5372
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 5371 GGGGGTCCGCTGTGTGTCGCCCGGGGCGACCGCTGGGCAATTTAGGCGCGCGGTGTGC 5312
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 5311 ACCCGTGGAGTGGCTTAAGCGGTGGACTTATCTCTGTGGAGAACCTTAGAGACCAACCATG 5252
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrClnVal 200
Db 5251 AGGTCCCGGGTTCACGGATAACTCTCTCCACAGTAGTGGCCCCAGAGTCTCCAGGTG 5192
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 5191 GCTCACCTCCATGCTCCACAGGAGCGGCGCAAGACCAACAGGTCGCCGCTGCATATGCA 5132

```

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 5131 GCTCAGGGCTATAAGGCTAGTACTCAACCCCTCTCTGCTGCAACACACCTGGGCTTTGGT 5072
QY 241 AlaTyrMetSerLysAlaHisGlyLeuAspProAsnIleArgThrGlyValArgThrIle 260
DB 5071 GCTTACATGTCGAAGCTCATGGGATCGATCCCTACATCAGGACCGGGGTGAGAACATTT 5012
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
DB 5011 ACCACTGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCTCTGGCAGCGGGGTGC 4952
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 4951 TCGGGGGGGCGGTATGACATAAATAATTTGTGACAGGTGCCACTCCACGGATGCCACATCC 4892
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 4891 ATCTTGGCATCGGCACCTGCTCTTGACCAAGCAGACAGCTCGGGGGCGAGACTGGTTGTG 4832
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 4831 CTCGCCACCGCCACCCCTCGGGGCTCGTCACTGTGCCCCATCCCAACATCGAGGAGGTT 4772
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 4771 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTAATC 4712
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 4711 AAGGGGGGGAGACATCTCATCTTCTGTCTTCAAGGAAGAGTCCGACGAACCTCGCGCA 4652
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 4651 AAGCTGTGCGATTGGGATCAATCGCTGGCTTACTACCGCTCTTGACGTGCTGCTC 4592
QY 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 4591 ATCCGACCGACGCGCGATGTTGCTGCTGGCAACCGATCCCTCATGACCGGCTATACC 4532
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
DB 4531 GCGGACTTCGACTCGGTGATAGACTGAATACGTTGCTGCCACCGACAGTCGATTTCAGC 4472
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 4471 CTTCACCTTACCTTCAACATGACAAATCACGCTCCCCCAGGATGCTGCTCCCCGACT 4412
QY 461 GlnArgArgGlyArgThrGlyArgGlyProGlyIleTyrArgPheValAlaProGly 480
DB 4411 CAACGTCGGGCGAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTGTGGCACCAGGG 4352
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 4351 GAGCGCCCTCCCGCATGTCGACTGCTCGTCTCTGAGAGTCTATGACGAGGCTGT 4292
QY 501 AlaTyrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 4291 GCTTGGTATGACTCAGCGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4232
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrPyrGluGlyValPheThrGlyLeu 540
DB 4231 CCGGGGCTTCCCGTGGCAGGACCATCTTGAAATTTTGGAGGGCGTCTTTACAGGCTTC 4172
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 4171 ACTCATATAGATGCCCATCTTCTATCCACAAAGCAGAGTGGGAGAACCTTCTCTTAC 4112
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
DB 4111 CTGTAGCGTACCAAGGACCGGTGGCTAGGCTCAAGCCCTCCCTCCATCGTGGGAC 4052
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600

DB 4051 CAGATGTGGAAGTGTATTGATTCGCTCAAGCCCAACCTTCCATGGGCCAACACCCTGCTA 3992
QY 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 3991 TACGACTGGGGCTGTTTTCAGATGAATCAATCACCTGACGACCCAGTCACCAATACATC 3932
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrThrValLeuValGlyGly 640
DB 3931 ATGACATGATGTTCGGCGGACCTGGAGTCTGACGACACCTGGGTCTCGTTGGCGGC 3872
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 3871 GTCTTGGCTGTTTGGCGCGGTATTGCTGTCACAGGCTGCGTGCATAGTGGCGAG 3812
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 3811 GTCTGCTTGTTCGGGAGCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 3752
QY 681 AspGluMetGluGluCys 686
DB 3751 GATGATGGAAGAGTGC 3734
RESULT 12
US-09-881-239-2
; Sequence 2, Application US/09881239
; Patent No. 6630298
; GENERAL INFORMATION:
; APPLICANT: CHEN, David Y.
; APPLICANT: ARCANSE, Phillip
; APPLICANT: TANDESE, Laura
; APPLICANT: GEORGE-NASCIMENTO, Carlos
; APPLICANT: COIT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: HCV ANTIGEN/ANTIBODY COMBINATION ASSAY
; FILE REFERENCE: 2302-16073 / PP16073.003
; CURRENT APPLICATION NUMBER: US/09/881,239
; CURRENT FILING DATE: 2001-06-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2058
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: representative NS3/4a conformational antigen
; NAME/KEY: CDS
; LOCATION: (1) .. (2058)
US-09-881-239-2
Alignment Scores:
Pred. No.: 0 Length: 2058
Score: 3565.00 Matches: 671
Percent Similarity: 99.13% Conservative: 9
Best Local Similarity: 97.81% Mismatches: 6
Query Match: 98.54% Indels: 0
DB: 4 Gaps: 0
US-09-930-591-2 (1-686) x US-09-881-239-2 (1-2058)
QY 1 MetaLapProfileThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1 ATGGCGCCCATCAGCGGTACGCCAGACAGACAGGGGCTCTTAGGGTGATATATAC 60
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 61 AGCTTAACCTGCGCGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCACTGCT 120
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
DB 121 GCCCAACCTTCTTCGGCAACGTGATCATATGGGGTGTCTGGACTGTCTACCGCGGGCC 180

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrIleValAsp 80
DB 181 GGAACGAGGACCATCGCGTCACCAAGGGCTCTGTCATCCAGATGATACCAATGATAGAC 240
QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
DB 241 CAAGACCTTGGCTGGCCGCTCCGCAAGGTAGCGGATCATTCACACCCCTGCATTCG 300
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
DB 301 GGTCTCTCGGACCTTTACCTGGTTCAGGACGACGCGGATGTCATTCGCGTCCGCGCG 360
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 361 GGTGATAGCAGGGGACGCTGCTCGCCCGGCCCATTTCTTACITGAAGGCTCTCTCG 420
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 421 GGGGGTCCGCTGTGTGTCGCGGGGACGCGGTGGCATATTTAGGGCCGCGGTGTC 480
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 481 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTAGAGACAACCATG 540
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
DB 541 AGGTCCCGCGTGTTCACGGATAACTCTCTCCACCATAGTGTGCGCCCGAGAGCTTCCAGGTG 600
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
DB 601 GCTCACTCCATGCTCCACAGCAGCGGCAAGAACCCAGGTCGCGGTGCTATATGCA 660
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
DB 661 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTCTGCTGCAACACTGGGCTTTGGT 720
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
DB 721 GCTTACATGTCCAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAT 780
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyLysCys 280
DB 781 ACCACTGGCAGCCCATCAGTACTCCACCTACGGCAAGTTCCTTCCCGCAGCGCGGTGC 840
QY 281 SerGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
DB 841 TCGGGGGCGCTTATGACATAATAATTGTGAGAGTGGCCACTCCACGGATGCCACATCC 900
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
DB 901 ATCTTGGGCATTGGCACTGCTTGACCAAGCAGAGACTCGGGGGGAGACTGGTTGTG 960
QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
DB 961 CTGCGCACCCGCCCTCCGCGGCTCGCTCACTGTGCCCATCCCAACATCGAGGAGTT 1020
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
DB 1021 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 1080
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
DB 1081 AAGGGGGGAGACATCTCTCTCTGTCATTCAAAGAAGAAGTGGCAGCAACTCGCGCA 1140
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
DB 1141 AAGCTGTGCGCATTTGGGCATCAATCGCGTCTACTACCGCGGTCTTGACGTGTCCGTC 1200
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
DB 1201 ATCCCGCCCATCGCGGATGTGTGCTGTGGCAACCATGCCCTCATGACCGGCTATACC 1260
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

DB 1261 GCGCACTTCGACTCGTGATAGACTGCATATCTGTGTCTACCCAGACAGTGCATTTTCAGC 1320
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
DB 1321 CTTGACCCCTACCTTCACCATTCAGACATACGCTCCCGCAAGATGCTGTCTCCGCACT 1380
QY 461 GlnArgAspGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
DB 1381 CAACTCGGGGACAGACTCGCGGGGAGGCGAGCATCTACAGATTTGTGGCCCGGG 1440
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
DB 1441 GAGCGCCCTCCGCGCATGTCGACTCGCTCTCTGTGAGTGTATGACGACGAGCTGT 1500
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
DB 1501 GCTTGGTATAGCTCACCGCCCGGAGACTACAGTTAGGTACGAGCGGTACATGACACC 1560
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
DB 1561 CCGGGGCTTCCGCTGTGCGAGCAACCATCTTGAATTTGGGAGGGCGCTCTTTACAGCGCTC 1620
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
DB 1621 ACTCATATAGATGCCACTTTCTATCCAGACAAAGAGAGTGGGGAGAACCTTCCTTAC 1680
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
DB 1681 CTGGTAGGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCCCATCGTGGGAC 1740
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
DB 1741 CAGATGTGGAAGTGTGTGATTCGCTCAAGCCCACTCCATGGGCCCAACCCCTGCTA 1800
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
DB 1801 TACAGACTGGGCGCTGTTCAGATGAAATACCTTCAGCGACCCAGTGCACCAATATATC 1860
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
DB 1861 ATGACATGCATGTGCGCGACCTGGAGTGTGTCACGAGCACTGGGTGCTCGTGGCGGC 1920
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
DB 1921 GTCCCTGGCTGTCTGGCGCGTATTCCTGTCAACAGGCTGCGTGTGCTATAGTGGCAGG 1980
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
DB 1981 GTCTGCTTGTCCGGGAGCGGCAATCATCTGACAGGAGTCTCTACCGAGAGTTC 2040
QY 681 AspGluMetGluGluCys 686
DB 2041 GATGAGATGGAAGATGC 2058

RESULT 13
US-09-881-654-1
; Sequence 1, Application US/09881654
; Patent No. 6632601
; GENERAL INFORMATION:
; APPLICANT: CHIEN, David Y.
; APPLICANT: ARCANDEL, Phillip
; APPLICANT: TANDESKS, Laura
; APPLICANT: GEORGE-NASCIEMTO, Carlos
; APPLICANT: COLT, Doris
; APPLICANT: MEDINA-SELBY, Angelica
; TITLE OF INVENTION: IMMUNOASSAYS FOR ANTI-HCV ANTIBODIES
; FILE REFERENCE: 2302-17039 / PP17039.002
; CURRENT APPLICATION NUMBER: US/09/881,654
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: 60/212,082
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/280,811

; PRIOR FILING DATE: 2001-04-02
 ; PRIOR APPLICATION NUMBER: 607280,867
 ; PRIOR FILING DATE: 2001-04-02
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2058
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:
 ; OTHER INFORMATION: representative NS3/4a conformational antigen
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2058)
 ; US-09-881-654-1

Alignment Scores:
 Pred. No.: 0 Length: 2058
 Score: 3565.00 Matches: 671
 Percent Similarity: 99.13% Conservative: 9
 Best Local Similarity: 97.81% Mismatches: 6
 Query Match: 98.54% Indels: 0
 DB: 4 Gaps: 0

US-09-930-591-2 (1-686) x US-09-881-654-1 (1-2058)

QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1 ATGGCGCCCATCAGCGGTACGCCAGCAGACAGAGGGGCCCTCTAGGGTGCAATACACC 60
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGluValGlnIleValSerThrAla 40
 DB 61 AGCCTAATGTCGCGGACAAACCAAGTGGAGGGTCCAGATTGTGCAACTGCT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 DB 121 GCCCAACCTTCCTGGCAACGTGATCAATGGGGTGTCTGGAGCTGTACACACGGGGCC 180
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 181 GGAAACGAGGACCATCGCGTCAACCAAGGTCCTGTATCAGATGTATACCAATGTAGAC 240
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
 DB 241 CAAGACCTTGTGGCTGGCGCTGCCAGAGGTAGCGGATCATTCACCCCTGCATCTGC 300
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 DB 301 GGCTCCTCGGACCTTTACCTGGTCAAGGACGCGCGATGTCATTCCTGCGCGCGGG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 361 GGTATACAGGGGACGCGCTGTCTCCCGCGGCCATTTCTACTTGAAGGCTCCTCG 420
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 421 GGGGGTCCGCTGTGTCCCGCGGGGCAAGCGGTGGGATATTTAGGGCGCGGTGTGC 480
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 481 ACCCGTGGAGTGGGTAAGGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGAACCATG 540
 QY 181 ArgSerProValPheSerAspAsnSerProProAlaValProGlnSerTyrGlnVal 200
 DB 541 AGGTCCCGGTGTTCAGGATACTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 600
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaLysAla 220
 DB 601 GCTCACCTCCATGCTCCCAAGGCGGCAAGACCAAGTCCCGGTGCGATATGCA 660
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAspProSerValAlaAlaThrMetGlyPheGly 240
 DB 661 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTCTGCAACACTGGGCTTTGGT 720

QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 721 GCTTACATGTCAGAGCTCATGGGATCATCTTACATCAGAGCCGGGTGAGAACAAAT 780
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
 DB 781 ACCACTGGCAGCCCATCATCGTACTCCACCTACGCAAGTTCCTTGGCAGCGGGTGC 840
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 841 TCGGGGGCGCTTATGACATAATAATTTGTACGAGTCCCATCCACGATGCCACATCC 900
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 901 ATCTTGGGCATGGCAGCTGTCTTGACCAAGCAGAGACTGGGGGGCAGACTGGTGTG 960
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 DB 961 CTTCGCCACCGCCACCCCTCCGGGTCCGTCACTGTGCCCATCCCAACATCGAGAGGTT 1020
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 DB 1021 GCTCTGTCCACCACCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCTCGAAGTATC 1080
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 DB 1081 AAGGGGGGAGACATCTCATCTCTCATTCAAAGAAAGTGCAGCAACTCCCGCA 1140
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 1141 AAGCTGGTGGCATTTGGGCATCAATGCCGTGCCCTACTACCGCGTCTTTGAGTGTCCGTC 1200
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 1201 ATCCCGCCCATCGGCGATGTTGTCGTGGCAACCGATGCCCTCATCGCGCTATACG 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 1261 GCGCACTTCGATCGGTGATAGCTGCAATACGTGTGTACCCAGAGATGCTGTCTCCGCACT 1320
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 1321 CTTGACCTCATCTTCACCATTTGAGACATCACGCTCCCCAGATGCTGTCTCCGCACT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 DB 1381 CAACCTCGGGCAGGACTGGCAGGGGGAAGCAGGCACTACAGATTTGTGGCACCGGGG 1440
 QY 481 GluArgProSerGlyWetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 DB 1441 GAGCGCCCTCTCGGCATGTTGCACTCGCTCTCTCTGTGAGTGTCTATGACGAGGTGT 1500
 QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 1501 GCTTGGTATGAGTCAACCGCCGAGACTACAGTAGGCTTACGAGCGTACATCAACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 DB 1561 CCGGGGCTTCCTCGGTGCGCAGGACATCTTGAATTTGGGAGGGCGCTTTTACAGGCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 DB 1621 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 1680
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrAsp 580
 DB 1681 CTGTGTAGCGTACCAAGCCACCGCTGTGGCTAGGCTCAAGCCCTCCCTCCATCTGGGAC 1740
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 1741 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCTCCATCGGCCCAACACCCCTCTA 1800

```

Qy 601 TyArgLeuGlyAlaValGlnAenGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TACAGACTGGCGCTCTTCAAGATGAATCAACCTGACGACCCAGTCAACAAATACATC 1860

Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGATGTGCGCCGACCTGGAGTCTGTCACGAGCACCTGGGTGCTCGTTGGCGGC 1920

Qy 641 ValLeuAlaLeuAlaValThrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTCTGCTCTTGGCGCGGTATTGCTGCTCAACAGGCTGCTGCTAGTGGGCAGG 1980

Qy 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 GTGCTCTTGTCCGGGAAGCGGCAATCATACCTGACAGGGAAGTCTCTACCGAGAGTTC 2040

Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAGATGAAGAGTGC 2058

RESULT 14
US-08-444-818-137
; Sequence 137, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hardin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8987 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..8985
US-08-444-818-137

```

```

Alignment Scores:
Pred. No.: 0
Score: 3565.00
Percent Similarity: 99.27%
Best Local Similarity: 97.81%
Query Match: 98.54%
Length: 8987
Matches: 671
Conservative: 10
Mismatch: 5
Indels: 0

```

```

DB: 3 Gaps: 0
US-09-930-591-2 (1-686) x US-08-444-818-137 (1-8987)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 3076 CTGGCGCCCATACAGCGCGTACGCCAGCAGCAAGAGGGGCTCTCTAGGTGATATACACC 3135

Qy 21 SerLeuThrGlyArgAspIleValGlnGluValGlnIleValSerThrAla 40
Db 3136 AGCTTAATCTGGCCGGGCAAAAAACAGTGAGGGTGAGGTCCAGATTGTGTCAACTGCT 3195

Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyVala 60
Db 3196 GCCCAAACTTCTCTGGCAACGTGCATCAATGGGTGTGTCTGGACTGTCTACACGGGCCC 3255

Qy 61 GlyThrArgThrIleAlaSerProIysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 3256 GGACGAGGACCATCGCTCACCAAGGGTCTCTCATCCAGATGTATACCAATGTAGAC 3315

Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 3316 CAGACCTTGTGGGTGGCTGGCTCCGCAAGGTAGCGCTCATTCACACCCCTGCACCTTGC 3375

Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 3376 GGCTCTCGGACCTTTACTGTGTACGAGGACGCGCATGTCTCCGTGCGCGCGG 3435

Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 3436 GGTGATAGCAGGGGCGAGCTGTGTGCGCCCGGCCCATTTCTACTTGAAGAGCTCTCTCG 3495

Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3496 GGGGTCTCGTGTGTGCCCCGCGGCGCACGCGGTGGGCATATTTAGGCGCGCGGTGTC 3555

Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3556 ACCCGTGGAGTGGCTAAGCGGTGACTTTATCTCTGTGGAGAACTGAGACACACCATG 3615

Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3615 AGGTCCCCGGTGTTCACGGATAACTCTCTCCACAGTAGTGTGCCCGCAGAGCTTCCAGGTG 3675

Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaThrMetGlyPheGly 240
Db 3676 GGTCACTCTCATGCTCCACAGCAGCGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA 3735

Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 260
Db 3736 GGTCAAGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGTCACACTGGGCTTTGGT 3795

Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 280
Db 3796 GCTTACATGTCCAGGCTCATGGGATCGATCTTAACATCAGGACGCGGGGTGAGAACAAAT 3855

Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 3856 ACCACTGGCAGCCCATCATCAGTACTCCACTACGGCAAGTTCTTTCGCCACGCGGGTGC 3915

Qy 281 SerGlyGlyValAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 3916 TCGGGGGGCGCTTATGACATAATAATTTGTGACGAGTGCCTCCACGATGCCACATCC 3975

Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 3976 ATCTTGGGCATCGGCATGTCTCTTCAAGCAGAGACTGCGGGGCGGAGACTGGTTGTG 4035

Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal 340
Db 4036 CTCGCCACCGCCACCCCTCCGGGCTCCGTCTCTGTGCCCCCATCCCAACATCAGGAGGTT 4095

Qy 341 AlaLeuSerThrThrGlyLysProPheTyrGlyLysAlaIleProLeuGluAlaIle 360

```

Db 4096 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGTATATCCCTTCGAGTAATC 4155
Qy 361 LysGlyGlyArgHisLeuPheCysHisSerLysLysLysCysAspGluLeuAlaAala 380
Db 4156 AAGGGGGGGGACATCTCATCTTCTGCAATCAAGAGAGAGTCCGACGAACTCGCGCA 4215
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4216 AAGCTGTGTCGATGGGCATCAATGCGGTGCTACTACCGCGCTTTCAGCTGTCGTC 4275
Qy 401 IleProThrSerGlyAspValValAlaValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4276 ATCCCGACACCGCGGATGTGTGCTGCTGCGCAACCGATCCCTCATGACCGGTATACC 4335
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4336 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTTCAGC 4395
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4396 CTTGACCTTACCTTACCATTTAGACCAATCACGCTCCCGGAGTGTCTCTCCGCACT 4455
Qy 461 GlnArgArgGlyArgThrGlyArgGlyValProGlyIleTyrArgPheValAlaProGly 480
Db 4456 CAACGTGGGGCAGACTGGCAGGGGAGCCAGGCATCAACAGATTGTGGCACCAGGG 4515
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4516 GAGCGCCCTCCCGCATGTCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4575
Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4576 GTTGTGTAGACTCAGCGCCGCGGACACTACAGTTAGGCTAGCGGTACATGAACACC 4635
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 4636 CCGGGCTTCCCGTGTCCAGGACCATTTGAATTTGGAGGGCGCTTTACAGGCTC 4695
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 4696 ACTCATATAGATGCCCATTTCTATCCAGCAAGCAGAGTGGGAGAACCTTCTCTTAC 4755
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 4756 CTGCTAGCTACCAAGCAGCGGTGCTGCTAGGCTCAAGCCCTCCCTCCCTCGTGGAC 4815
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 4816 CAGATGTGGAGTGTGTTGATTGCGCTCAAGCCACCCCTCCATGGGCAACACCCCTGCTA 4875
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 4876 TACAGCTGGCGCTGTTCAATGAATCAATCACCCTCAGCGACCCCACTCAACAAATATC 4935
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 4936 ATGACATGCTATGCGGCGACCTGGAGTGTGTACGAGCACCTGGGTGTGTGGCGC 4995
Qy 641 ValLeuAlaLeuAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 4996 GTCCGTGCTGCTTGGCGGATGCTGCTCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCT 5055
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5056 GTCCGTGCTGCTTGGCGGAGCCGCAATCATACCTGACAGGAGTCTCTTACCGAGGTT 5115
Qy 681 AspGluMetGluGluCys 686
Db 5116 GATGAGATGGAAGTGC 5133

RESULT 15

US-08-444-818-53

; Sequence 53, Application US/08444818
; Patent No. 6150087
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANBV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,818
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5360 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..5360
; US-08-444-818-53

Alignment Scores:
Pred. No.: 0 Length: 5360
Score: 3555.00 Matches: 668
Percent Similarity: 99.42% Conservative: 13
Best Local Similarity: 97.52% Mismatches: 4
Query Match: 98.26% Indels: 0
DB: 3 Gaps: 0

US-09-930-591-2 (1-686) x US-08-444-818-53 (1-5360)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 930 CTGGCGCCCATCATCGCGGTAGCGCCACGACACCAAGGGCCCTCTAGGTGCATAATCACC 989
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 990 AGCTTAATCTGCGCGGCAAAACCAACCAAGTGGAGGTGAGTCCAGATTGTGTCACTGCT 1049
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db 1050 GCCCAACCTTCTTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAAGGGGCC 1109
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1110 GGAACGAGGACCATCGCGTCACCAAGGTCCTGTCTATCCAGATGTATACCAATGTAGAC 1169
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100

Db 1170 CAAGACCTTGTGGGCTGGCCGCTCGGCAAGGATGACCGCTCATTGACACCCCTGCATTCG 1229
Qy GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 1230 GGCTCTCGGACCTTTACCTGTGTACAGGACGCGCATGTCTATTCCTGCGCGCGG 1289
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 1290 GGTGATAGCAGGGGAGCTGCTGTGCGCCCGGCGCCATTTCTCTGAAAGGCTCTCTCG 1349
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaValCys 160
Db 1350 GGCGGCTCGCTGTGTGCGCGCGGCGACGCGGTGGGCATATTTAGGCGCGGTGTGC 1409
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
Db 1410 ACCCGTGGAGTGGCTAAGCGGTGGATTTATCTCTGTGGAGAACCCTAGAGCAACCATG 1469
Qy 181 ArgSerProValPheSerAspAenSerSerProAlaValProGlnSerTyrGlnVal 200
Db 1470 AGGTCCCGGTGTTCAGGATTAATCTCTCCACCAAGTAGTCCCGCAGAGCTCCAGGTG 1529
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaIleTyrAla 220
Db 1530 GCTCACTCTCCATGCTCCACAGCAGCGGCAAGCAACCAAGGTCCTCCGCTGCATATGCA 1589
Qy 221 AlaGlnTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 1590 GCTCAGGCTATAGGTGCTAGTACTCAACCCCTCTGTGTGTCACACATGGGCTTTGGT 1649
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 1650 GCATTACATGTCGAAGGCTCATGGATCATCTTAACATCAGGACCGGGGTGAGAACAT 1709
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIlePheLeuAlaAspGlyLysCys 280
Db 1710 ACCACTGGCAGCCCATCATGCTACTCCACTCAGCAAGTTCTTCCGACGCGGGTGC 1769
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 1770 TCGGGGGCGCTTATGACATAATAATTTGTGACGAGTGGCCACTCCACGGATGCCACATCC 1829
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 1830 ATCTTGGGCATCGGCATGCTCTTGACCAAGCAGAGACTCGGGGGCGAGACTGGTGTG 1889
Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 1890 CTGCCACCGCCACCCCTCCGGGCTCCGTCACTGTGCCCATCCACATCCGAGAGGT 1949
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1950 GCTCTGCCACACCGGAGAGATCCCTTTTAAGCAAGGCTATCCCGCTCGAAGTAATC 2009
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2010 AAGGGGGGAGACATCATCTCTGTCAATTAAGGAAGTGCAGCAACTCGCGCA 2069
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2070 AAGTGTGTGGTGGCATCAAGTGGGCTGCTTACCGCGTCTTGAGCTGTCCGTC 2129
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2130 ATCCCGACACGCGGATGTTGTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACC 2189
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2190 GCGGACTTCGACTCGGTGATGAGACTGCAATACGTGTGTACCCACAGACAGTCAATTCAGC 2249
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 2250 CTTGACCCCTACCTTCAACATTTGAGACAAATACGCTCCCCCAGGATGCTGTCTCCCGCACT 2309

Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 2310 CAACGTGGGGCAGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGCACCAGGG 2369
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 2370 GAGCGCCCTCCGCGCATGTTCCGACTCGTCGCTCTGTGAGTGCTATGACGAGGCTGT 2429
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 2430 GCTTGGTATGAGCTACGCCCGCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 2489
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 2490 CCGGGCTTCCCGTGTGCCAGACCATCTTCAATTTGGAGGGCGCTTTACAGGCGCTC 2549
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProThr 560
Db 2550 ACTCATATAGATGCCCATTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCTTAC 2609
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 2610 CTGGTAGGTACCAAGCCACCGTGTGGCTAGGCTCAAGCCCTCCCGCATCGTGGAC 2669
Qy 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 2670 CAGATGTGGAAGTGTGTTGATTGCGCTCAAGCCCATCTCCATGGGCCAACACCCCTGCTA 2729
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 2730 TACAGACTGGGCGCTGTTCAAGATGAATCACTGACGACCCAGCAGTCAACCAATACATC 2789
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 2790 ATGACATGCATGTCCGCGCACCTGGAGGTCTGCACGACACCTGGGTGCTCGTTGGCGC 2849
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 2850 GTCTGTCTGTCTTGGCGCGTATTGCTGTCAACAGGCTCGTGTGCTATAGTGGCAGG 2909
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 2910 GTCTGTCTGTCCGGGAAGCCGCAATCATCTGACAGGAAGTCTCTACCGAGAGTTC 2969
Qy 681 AspGluMetGluGlu 685
Db 2970 GATGAGATGGAAG 2984

Search completed: September 16, 2004, 00:45:35
Job time : 330 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:57 ; Search time 712 Seconds
(without alignments)
4093.067 Million cell updates/sec

Title: US-09-930-591-2

Perfect score: 3618

Sequence: 1 MAPITAYAAQTRGLGCIIT.....PALIPDREVLVREFDEMEEC 686

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framep_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US0930591/runat_15092004.164645.7086/app_query.fasta_1.839
-DB=N_Geneseq_29Jan04 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOCPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0930591 @CGN 1.1 470 @runat_15092004.164645.7086 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3618	100.0	2061	6 AAD34500	Aad34500 Hepatitis
2	3618	100.0	2061	6 AAD31767	Aad31767 Hepatitis
3	3618	100.0	2061	9 AAD60868	Aad60868 Hepatitis
4	3574	98.8	7310	1 AAN92106	Aan92106 Combined
5	3574	98.8	7310	1 AAN90336	Aan90336 Composite
6	3574	98.8	7310	2 AAG98221	Aag98221 Hepatitis
7	3574	98.8	8316	3 AAA5296	Aaa5296 cDNA sequ
8	3574	98.8	9133	2 AAZ07656	Aaz07656 Nucleotid

9	3574	98.8	9185	2	AAQ10566	Aaq10566 Hepatitis
10	3574	98.8	9185	3	AAA75297	Aaa75297 Sense str
11	3574	98.8	9401	2	AAT12710	Aat12710 Hepatitis
12	3574	98.8	9401	2	AAT99981	Aat99981 HCV poly
13	3574	98.8	9401	2	AAV09989	Aav09989 HCV poly
14	3574	98.8	9401	6	AAD35043	Aad35043 Hepatitis
15	3571	98.7	6905	1	AAN92103	Aan92103 Combined
16	3571	98.7	9185	2	AAZ26737	Aaz26737 Nucleotid
17	3570	98.7	9185	2	AAQ05956	Aaq05956 Sense str
18	3568	98.6	6299	4	AAF83669	Aaf83669 HCV NS3/4A
19	3568	98.6	9185	2	AAQ00459	Aaq00459 Hepatitis
20	3567	98.6	8316	2	AAQ05955	Aaq05955 Hepatitis
21	3565	98.5	2058	6	ABK15344	Abk15344 Hepatitis
22	3565	98.5	2058	6	AAD29795	Aad29795 HCV-1 NS3
23	3565	98.5	2058	7	ABX14410	Abx14410 DNA encod
24	3565	98.5	2058	9	ADC06768	Adc06768 HCV mutan
25	3565	98.5	5400	2	AAQ21744	Aaq21744 Compiled
26	3555	98.3	5360	1	AAN90327	Aan90327 Hepatitis
27	3552	98.2	5300	1	AAV92097	Aav92097 Hepatitis
28	3551	98.1	9646	2	AAV59361	Aav59361 Hepatitis
29	3551	98.1	9646	6	ABK87285	Abk87285 cDNA enco
30	3551	98.1	9646	7	ACA62466	Aca62466 HCV H77 C
31	3551	98.1	12980	2	AAV59364	Aav59364 Hepatitis
32	3551	98.1	12980	6	ABK87286	Abk87286 Hepatitis
33	3551	98.1	12980	7	ACA62469	Aca62469 DNA encod
34	3551	98.1	16622	3	AAZ36212	Aaz36212 Nucleotid
35	3550	98.1	9502	2	AAQ74770	Aaq74770 Hepatitis
36	3544	98.0	9518	5	AAQ03778	Aaq03778 Hepatitis
37	3544	98.0	9599	2	AAZ24833	Aaz24833 Infectiou
38	3544	98.0	9599	2	AAZ24832	Aaz24832 Infectiou
39	3544	98.0	9599	4	AAZ86938	Aaz86938 Nucleotid
40	3544	98.0	9611	5	AAZ86646	Aaz86646 Nucleotid
41	3544	98.0	9611	5	AAZ86645	Aaz86645 Nucleotid
42	3544	98.0	9611	5	AAZ86647	Aaz86647 Nucleotid
43	3544	98.0	9611	5	AAZ86648	Aaz86648 Nucleotid
44	3544	98.0	10803	7	ABX10617	Abx10617 MKO-Z nuc
45	3543	97.9	9379	2	AAQ36209	Aaq36209 Composite

ALIGNMENTS

RESULT 1
AAD34500
ID AAD34500 standard; DNA; 2061 bp.
XX
AC AAD34500;
XX
DT 16-JUL-2002 (first entry)
XX
DE Hepatitis C virus NS3/4A protein encoding DNA.
XX
KW Hepatitis C virus; HCV; NS3/4A protein; therapy; HCV infection; vaccine;
KW virucide; gene; ds.
XX
OS Hepatitis C virus.
XX
PH Key Location/Qualifiers
FT CDS 1..2061
FT /*tag= a
FT /product= "HCV NS3/4A protein"
XX
PN WO200214362-A2.
XX
PD 21-FEB-2002.
XX
PF 15-AUG-2001; 2001WO-IB001774.
XX
PR 17-AUG-2000; 2000US-0225767P.
PR 29-AUG-2000; 2000US-0229175P.
PR 03-NOV-2000; 2000US-00705547.
XX
PA (TRIP-) TRIPEP AB.
XX

PI Sallberg M;
 DR WPI: 2002-339446/37.
 DR P-PSDB; AAE21837.
 XX
 PT Novel hepatitis C virus NS3/4A peptide useful for diagnosing presence or
 PT absence of hepatitis C virus in a subject and for preparing a medicament
 PT for treating hepatitis C virus infection.
 XX
 PS Claim 1; Page 64-65; 90pp; English.
 XX
 CC The present invention relates to novel hepatitis C virus (HCV) NS3/4A
 CC proteins and their corresponding polynucleotides. NS3/4A sequences are
 CC useful for identifying the presence or absence HCV in a subject. They are
 CC useful for preparing a medicament used for treating or preventing HCV
 CC infection. Sequences of the invention are also used as vaccines. The
 CC present sequence is a DNA encoding HCV NS3/4A protein
 XX
 SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,44e-245 Length: 2061
 Score: 2618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x AAD34500 (1-2061)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1 ATGGCGGCTATACGGGCTATGCCAGACAGAGGGGGCTTTTGGGATGCATAATCACC 60
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 61 AGCTTGACCGCGGGGACAAACACAGGTGGAGGGTGAGGTTGAGATCGTGTCAGTGGT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 DB 121 GCCCAGACACTTCTTGGCAACCTGACCTAACCGGGGTGTGTGGAGCTGTCTACCATGGAGCC 180
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 181 GGAAACAGGACCATGTGGTACCTAACGGGGTGTGTGGAGCTGTCTACCATGGAGCC 240
 QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
 DB 241 CAAGACCTCGTAGGCTGGCGGCTCCCAAGGTGGCGGCTCATTAACCATGCACTTGC 300
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 301 GGCTCCTCGGACCTTACCTGGTACAGGACGCGCGATGCTATCTCTGAGGCGCGACGG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 361 GGTGATGGCAGGGGACGCTGCTTTCGCCCGGCTATCTCTACTTGAAGGCTCTCTCG 420
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 421 GGAGGCGCTCTGCTGTGCCCGGAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC 480
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 481 ACCCGTGGAGTGGCTAAGGGGGTGGACTTTCATCCCGGTAGAGAGCTTTAGAGCAACATG 540
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 541 AGGTCCCGCGGTGTTCTCAGACAACCTCTCCCAACAGCAGTGGCGGCGGAGCTACCAAGTG 600
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 601 GCCCACCCTGCATGCTCCACCGGAGGGGTAGAGCACCAAGGTCCCGCGGATACGCCA 560

QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheCly 240
 DB 661 GGTCAAGGCTACAGGTGTGGTGTCAACCCCTCGTGTGTCACCAATGGGCTTTGGT 720
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleAlaThrGlyValArgThrIle 260
 DB 721 GCTTACATGTCCAAAGCCCATGGGATTGATCTTAACATCAGGACTGGGTGGAGCAATT 780
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
 DB 781 ACTACTGGCAGCCGATCAGCTATTCACCTACGCAAGATTCTTGGCAGCGGGGTGT 840
 QY 281 SerGlyGlyValAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 841 TCAGGGGGTGTCTATGACATAATAATTGTGACGAGTGCCTCCACGGATGCACATCC 900
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 DB 901 ATCTTGGGCAITGGCACTGTCTTGACCAAGCAGACCGCGGGGGGAGACTGACTGTG 960
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluVal 340
 DB 961 CTCGCCACCGTACCCCTCCGGGCTCCGTCACCTGTGCCCATCTTAACATCGAGAGGTT 1020
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 DB 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAAGCAAT 1080
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 DB 1081 AAGGGGGGGAGACATCTCATCTTGCACCTCAAGAAGAAAGTGCAGAGCTCCCGCA 1140
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 DB 1141 AAATGTGTGGTGGGGGTCAATGCCGTGCTTACTACCGCGGCTTGATGTCTCGTC 1200
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 DB 1201 ATCCCGACAGTGTGTGCTGTGTGCAACTGACGCCCTCATGACCGGCTTTTACC 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 DB 1261 GCGGACTTCGATTGGTGATAGACTGCACACGTGTACCCAGACAGTGCATTCAGC 1320
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 DB 1321 CTTGACCCCTACCTTCAACATGAGACAATCAGCTTCCCGAGGATGCTGTCTCCGTA 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyPheProGlyIleTyrArgPheValAlaProGly 480
 DB 1381 CAACGTCCGGGTAGACTGGCAGAGGGAAGCAGGCACTACACATTTTGGGACCGGG 1440
 QY 481 GlnArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 DB 1441 GAGCGTCTCTTGGCATGTTTGTGCTGTCTCTCTCGAGTGTCTATGACGGGGTGT 1500
 QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 DB 1501 GCTTGGTATGAGCTTACGCCCGCGAGACCACTAGGCTTAGGCTACGAGCATCATGAAC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
 DB 1561 CCGGAGCTTCCCGTGGCAAGACCATCTTGAATTTTGGAGGGCGCTCTTACGGGTCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 DB 1621 ACCCAGATAGACGGCCCATCTTCTATCCAGACAAGCAGAGTGGGAGAAACCTTCCCTAT 1680
 QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 DB 1681 CTGTAGGCTACCAAGCCACCGTGTGGTAGAGCTCAAGCCCTTCCCCGTGCTGGGAC 1740

QY 581 GlnMetTrpLysCysLeuLeuArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 DB 1741 CAGATGGGAAGTGTGATCCGCTCAAGCCACCCCTCATGGGCAACACCTCTGTCTA 1800
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 DB 1801 TATAGACTGGCGCTGTCCGAATGAAGTCAACCTCAGCCACCCAGTCAAGTATATC 1860
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 DB 1861 ATGACATGTATGTCGGCTGACCTGGAGTGGTCAACGAGTACCTGGGTGCTCGTTGGCGGC 1920
 QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 1921 GTTCTGGCTCTTTGGCCGGTATTCCTATCCACAGGCTGGTGTCTATAGTAGTAGG 1980
 QY 661 IleValLeuSerGlyLysProAlaIleLeuProAspArgGluValLeuTyrArgGluPhe 680
 DB 1981 ATTGTCTTGTCCGAAAGCCGGCAATCATACCCGACAGGGAATCTCTACCGGGAGTTC 2040
 QY 681 AspGluMetGluGluCys 686
 DB 2041 GATGAATGAAGAGTGC 2058
 RESULT 2
 AAD31767
 ID AAD31767 standard; DNA; 2061 BP.
 XX AAD31767;
 XX AC
 XX AC
 DT 18-JUN-2002 (first entry)
 XX DE
 XX DE Hepatitis C virus (HCV) NS3/4A DNA coding region.
 XX KW
 XX KW Hepatitis C virus; HCV infection; virucide; fungicide; antibiotic; antibacterial;
 XX KW cytosstatic; immunostimulant; vaccine; ribavirin; immune response; cancer;
 XX KW ds.
 XX OS
 XX OS Hepatitis C virus.
 FH Key Location/Qualifiers
 FT CDS 1..2061
 FT FT /*tag= a
 FT FT /product= "HCV NS3/4A protein"
 PN WO200213855-A2.
 XX 21-FEB-2002.
 XX 15-AUG-2001; 2001WO-IB001808.
 XX 17-AUG-2000; 2000US-0225767P.
 XX 29-AUG-2000; 2000US-0229175P.
 XX 03-NOV-2000; 2000US-00705547.
 XX (TRIP-) TRIPEP AB.
 XX Sallberg M, Hultgren C;
 XX WPI; 2002-241837/29.
 XX P-PSDB; AAE19900.
 XX Vaccine compositions for treating and preventing disease, preferably
 FT hepatitis C virus infection, comprises ribavirin and antigen that has
 FT epitope present in hepatitis C virus.
 XX
 XX Claim 1; Page 94-95; 120pp; English.
 XX The invention relates to a composition comprising ribavirin and an
 CC antigen preferably non structural 3 protein (NS3)/4A fragment of
 CC hepatitis C virus (HCV) genome or a peptide or nucleic acid of HCV
 CC sequence. The composition is useful for enhancing an immune response to a
 CC hepatitis C antigen in humans, domestic, sport or pet species and as

CC vaccines for treating and preventing HCV infections. The composition is
 CC also useful for treating viral, bacterial, fungal diseases and cancer.
 CC The present sequence is HCV NS3/4A DNA coding region
 XX Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,44e-245 Length: 2061
 Score: 3618.00 Matches: 696
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-930-591-2 (1-686) x AAD31767 (1-2061)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 1 ATGGCGCCTATACGGCCTATGCCAGCAGACAAGGGGCTTTTGGGATGATAATACCC 60
 QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 61 AGCITGACCGCGCGGACAAAACCAGGTGGAGGTGAGGTTCAGATCGTGTCACTGCT 120
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 DB 121 GCCCAGACTTCTTGGCAACCTGCATTAAACGGGTGTGTGGACTGTCTACCATGGAGCC 180
 QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 181 GGAAACAGGACCATTCGCTACCTAAAGGTCTCTGTATTCCAGATGTACACCAATGTGAC 240
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 241 CAAGACCTCTAGGTGGCGCGCTCCCAAGGTGCCCTCATTAACACCATGCACTTGC 300
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 DB 301 GGCTCCTCGGACCTTTTACCTGTGTACAGAGCAGCGCGATGTCAITCTCTGTGCGCGAGCG 360
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 DB 361 GGTGATGGCAGGGCAGCGCTGCTTTTGGCCCGCGCTATCTTCTTACTTGAAGGCTTCTCG 420
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 421 GGAGGCCCTCTGCTGTGCGCGCAGACATGCGGTAGGCATATTTCAGAGCGCGGTATGC 480
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 DB 481 ACCCGTGGAGTGGCTAAGCGGTGACTTCATCCCGGTAGAGAGCTTAGACAACACCATG 540
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 541 AGGTCCCCGGTGTCTTCAGACAACTCTCCCAAGGTGGAGTGGCGCGCATACCAAGT 600
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 601 GCGCACCTGTCATGCTCCACCGCGAGCGGTAAAGACCAAGGTCCCGCGCATACCGCA 660
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 661 GGTGAGGCTACAGGTGCTGTGTCTCAACCCCTCCGTTCGTGCAACATGGGCTTTGGT 720
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 721 GCTTACATGTCCAGGGCCCATGGGATTGATCTAAACATCAGGACCTGGGAGTGGAGCAATT 780
 QY 261 ThrThrGlySerProIleThrThrThrThrThrThrThrThrThrThrThrThrThrThr 280
 DB 781 ACTACTGGCAGCGCCGATACATGATTCACCTACCGCAAGTTCCTTTCGCGACGCGGGTGT 840
 QY 281 SerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSer 300

Db 841 TCAGGGGTGCTATGACATAATATTTGTGACGAGTGCCACTCCAGGATGCAACATCC 900
 QY 301 ILeuLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 901 ATCTTGCGCATTTGGCACTGTCTCTTGACCAAGACAGACCGCGGGGCGAGACTGACTGTG 960
 QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 961 CTCGCCACCGCTACCCCTCCGGGCTCGCTCACTGTGCCCCATCCATAACATCGAGAGGTT 1020
 QY 341 AlaLeuSerThrThrGlyIleProPheThrGlyIleProLeuGluAlaIle 360
 Db 1021 GCTCTGTCCATACCGGAGAGATCCCTTTATGGCAAGGCTATTCCCTTGAGCAATT 1080
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaAla 380
 Db 1081 AAGGGGGGAGACATCTCATCTCTGCCACTCAAGAAGAAGTCCGACGAGCTCGCCGCA 1140
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaIleThrValArgGlyLeuAspValSerVal 400
 Db 1141 AAACGTGTCGGGTGGGGGTCAATGCGGTGCTTACTACCGCGGCTTGATGTGTCGTC 1200
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 1201 ATCCGACCACTGTGAGGTTGTCTGTGCAACTGACGCCCTCATGACGGCTTTACC 1260
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 1261 GCGCACTTCGATTCGGTGATAGTCAACACAGTGTGTCAACCCAGACAGTCGACTTCAGC 1320
 QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 1321 CTTGACCTTACTTTCACCATTTGACATCATCGCTTCCAGGATGCTGTCTCCGTACT 1380
 QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleThrArgPheValAlaIleProGly 480
 Db 1381 CAACGTGCGGGTAGGACTGGCAGAGGGAAGCCAGGCACTTACAGATTTGTGGCACCGGG 1440
 QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysThrAspAlaGlyCys 500
 Db 1441 GAGGCTCTCTTGGCATGTTTGAATCTGCTGTCTCTGCGAGTGCTATGACGGGGTGT 1500
 QY 501 AlaThrPyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaIleMetAsnThr 520
 Db 1501 GCTTGTATGAGTTACGCCCGCGGACACACAGTTAGGCTACGAGCATACATGAACACC 1560
 QY 521 ProGlyLeuProValCysGlnAspHisGluGluPheThrGluGlyValPheThrGlyLeu 540
 Db 1561 CCGGGACTTCCCGTGTGCGCAAGACCATCTTGAATTTTGGAGGGCGGCTTTACGGGTCTC 1620
 QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 1621 ACCACATAGAGCCCATCTTCTATCCAGACAAAGCAGAGTGGGAAACCTTCCCTAT 1680
 QY 561 LeuValAlaIleThrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrTrpAsp 580
 Db 1681 CTGTAGCGTACCAAGGACCGGTGTGGGTAGAGCTCAAGCCCTCCCGCGTGTGGGAC 1740
 QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 1741 CAGATGTGGAGTGCTTGTATCCGCTCAAGCCACACCTCCATGGGCCAACACCTCTGCTA 1800
 QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 1801 TATAGACTGGCGGCTGTCCAGATGAAGTCACTCAGCAGCAGCCAGTCAACCAAGTATATC 1860
 QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 1861 ATGACATGTATGTCGGTGTGCTGGAGTGTGTACGAGTACCTGGGTGCTTGTGGCGC 1920
 QY 641 ValLeuAlaAlaLeuAlaIleThrCysLeuSerThrGlyCysValValIleValGlyArg 660

Db 1921 GTTGTGGTGCTTTGGCCCGCTATTGCTCTATCCACAGGCTCGTGGTCAATAGTAGGTAGG 1980
 QY 661 ILeuValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuThrArgGluPhe 680
 Db 1981 ATTGCTTTGTCCGAAAGCCGCAATCATACCCGACAGGAAGTCTCTACCGGGAGTTC 2040
 QY 681 AspGluMetGluGluCys 686
 Db 2041 GATGAATGGAAGAGTGC 2058
 RESULT 3
 ID AAD60868 standard; DNA; 2061 BP.
 AC AAD60868;
 DT 15-JAN-2004 (first entry)
 XX Hepatitis C virus NS3/4A DNA.
 XX Ribavirin; vaccine; immune response; infection; therapy; immunostimulant;
 KW virucide; ds.
 XX Hepatitis C virus.
 OS
 FH Key Location/Qualifiers
 FT CDS 1..2061
 FT /*tag= a
 FT /product= "Hepatitis C virus protein"
 XX US2002136740-A1.
 PD 26-SEP-2002.
 PF 15-AUG-2001; 2001US-00929955.
 PR 17-AUG-2000; 2000US-0225767P.
 PR 29-AUG-2000; 2000US-0229175P.
 XX (SALL/) SALLBERG M.
 PA (HULT/) HULTGREN C.
 XX Sallberg M, Hultgren C;
 DR WPI; 2003-764978/72.
 DR P-PSDB; ABW00351.
 XX Vaccine compositions for treating and preventing disease, preferably
 PT hepatitis C virus infection, comprises ribavirin and antigen that has
 PT epitope present in hepatitis C virus.
 XX Claim 1; Page 60-61; Opp; English.
 CC The invention relates to a composition comprising ribavirin and an
 CC antigen, where the antigen is derived from a hepatitis virus. The vaccine
 CC is useful in enhancing the immune response to a hepatitis C antigen where
 CC the composition is delivered to an animal identified as requiring an
 CC enhanced immune response. The vaccine is useful in the treatment and
 CC prevention of hepatitis C infection. The present sequence is Hepatitis C
 CC virus NS3/4A DNA
 XX
 SQ Sequence 2061 BP; 427 A; 616 C; 571 G; 447 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 2,44e-245 Length: 2061
 Score: 3618.00 Matches: 686
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 9
 US-09-930-591-2 (1-686) x AAD60868 (1-2061)

Qy 1 MetAlaProIleThrAlaIleAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1 ATGGCGCTATCATCGGCTATGCCAGCAGAACAGGGGCTTTTGGATGCATAATCAC 60
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 61 AGCTTGACCGCGCGGACAAACACAGGTGGAGGGTGAGGTTTCAGATCGTGCAACTGCT 120
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValThrHisGlyAla 60
Db 121 GCCCAGACTTCTTGGCAACCTGCATTAAACGGGGTGTGTGGACTGTCTACCATGGAGCC 180
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 181 GGRACAGAGCAATGGCTCAGCTTAGGGTCTGTATCCAGATGTACACCAATGTGGAC 240
Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 241 CAAGACCTGTAGCTGGCGGCTGCCAAGGTGCCCGCTCAATTAACACCATGCATCTGC 300
Qy 101 GlySerSerAspLeuThrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db 301 GGCTCCCTCGGACCTTTACCTGGTACAGGACAGCCGATGTCAITCTGTGCGCCGACGG 360
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 361 GGTGATGGCAGGGCAGCCTGCTTTCGCCCGGCTATCTCTTACTTGAAGGCTCCTCG 420
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 421 GGAAGGCCCTCTGTGTGCCCGCAGGACATGCCGTAGGCATATTCAGAGCGCGGTATGC 480
Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 481 ACCCGTGGAGTGGCTAAGCGGTGGACTTCATCCCCGTAGAGACTTAGAGACAACCATG 540
Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
Db 541 AGTCCCGGTGTCTCAGACAATCTCTCCCAACAGCAGTGCCTCCAGAGTACCAAGTG 600
Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 601 GCCCACTGTGATGCTCCACCGCGGCGGGTAAGACACCAAGTCCCGCGCATACGCA 660
Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 661 GCTCAGGGCTACAAGGTGCTGGTGTCAACCCCTCCCTGCTGCTCAACAATGGGCTTTGGT 720
Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db 721 GCTTACATGTCCAAGGCCCATGGGATGTATCTTAACATCAGGACTGGGGTGAGACAATT 780
Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 781 ACTACTGGCAGCCGATCGATATTCACCTAGGCAAGTTCCTTGGCCGACGGGGTGT 840
Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 841 TCAGGGGGTGTATGACATAATAATTTGTACAGAGTGCACATCCACCGATGCAACATCC 900
Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 901 ATCTTGGCATGGCATGTGCTTGTACACAGCAGAGACCGGGGGGGGAGACTGACTGTG 960
Qy 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 961 CTCGCCACCGTACCCCTCCGGGCTCGCTCACTGTGCCCATCTCTTAACATCGAGGAGTT 1020
Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 1021 GCTCTGTCCACTACCGGAGAGATCCCTTTTATGGCAAGGCTATTCCTCTTGAGCAATT 1080
Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380

Db 1081 AAGGGGGGAGACATCTCATCTTCTGCCACTCAAGAAAGAGTGGAGAGCTGCCGCA 1140
Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 1141 AAACCTGGTGGCTTGGCGCTCAATGCCCTGGCTTACTACCGCGGCTTGATGTGTCCGTC 1200
Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 1201 ATCCGACAGTGTGTGACGTGTGTGTGTGGCAACTGACGCCCTCATGACCGGTTTACC 1260
Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 1261 GSCGACTTCGATTCGTTGATAGACTCAACACGCTGTGTCAACCCAGACAGTGCAGTTACG 1320
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 1321 CTTGACCTTACCTTACCATTGAGACAATCAGCTTCCCCAGGATGCTGTCTCCGTTACT 1380
Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 1381 CAACGTCGGGTAGGACTGGCAGAGGAGGAGCAGCATCTACAGATTTGTGGCACCGGG 1440
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 1441 GAGCGTCCCTTCTGGCATGTTTACTCTGTCTGTCTCGAGTGTCTATGACCGGGTGT 1500
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 1501 GCTTGTATGAGCTTACGCCCGCGAGACACAGCTTAGGCTACGAGCATACATGAACACC 1560
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
Db 1561 CCGGACTTCCGTTGTGCAAGACCATCTTGAATTTTGGGAGGGCGTCTTTACGGGTCTC 1620
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 1621 ACCCAGTAGAGCCCACTTCTTATCCAGACAAGCAGAGTGGGGAACACCTTCCCTAT 1680
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 1681 CTGGTAGCGTACCAAGCCACCGTGTGCGCTAGAGCTCAAGCCCTCCCGCTCGTGGGAC 1740
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 1741 CAGATGTGGAAGTGTGTATCGCTCAAGCCCACTCCATGGGCCCAACACCTCTGCTA 1800
Qy 601 TyrArgLeuGlyValAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 1801 TATAGCTGGCGCTGTCCAGATGCAAGTCAACCTGACGCCACCCAGTCACCAAGTATATC 1860
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 1861 ATGACATGTATGTCCGCTGACCTGGAGTGTGTGACAGTACCTGGGTCTCTGTTGGCGGC 1920
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 1921 GTTCTGGCTGTCTTGGCGCGGTATTCCTATCCACAGAGTGGTGTCTAGTAGTAGTAG 1980
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 1981 ATTGTCTTGTCCGAAAGCCGCAATCATATCCGACGAGGAAGTCTCTACCGGGAGTTC 2040
Qy 681 AspGluMetGluGluCys 686
Db 2041 GATGAATGGAAGAGTGC 2058
RESULT 4
ID AAN92106
XX AAN92106 standard; DNA; 7310 BP.
AC AAN92106;
XX

DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 DE Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
 DE clones K9-1 through 15e.
 XX Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
 XX Hepatitis C virus.
 OS
 FH Key Location/Qualifiers
 FT CDS 3..7310
 FT /*tag= a
 XX
 PN EP182126-A.
 PD
 XX 31-MAY-1989.
 PF 18-NOV-1988; 88EP-00310922.
 XX
 PR 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.
 PR 26-FEB-1988; 88US-00161072.
 PR 26-MAY-1988; 88US-00191263.
 PR 26-OCT-1988; 88US-00263584.
 PR 14-NOV-1988; 88US-00271450.
 XX
 PA (CHIR) CHIRON CORP.
 PA (CHIR) CHIRON CORP.
 XX
 PI Houghton M, Choo QL, Kuo G;
 XX
 DR WPI; 1989-159274/22.
 DR P-PSDB; AAP92050.
 XX
 PT Purified hepatitis C virus - and associated nucleic acids and
 PT polypeptide(s).
 XX
 PS Claim 3; Fig 47-1 - 47-8; 139pp; English.
 XX
 CC It is a double-stranded nucleotide sequence of the open reading frame
 CC (ORF) (tag a) extending through clones K9-1 to 15e of hepatitis C virus
 CC (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
 CC detect the presence of HCV nucleic acids in samples. The polypeptide(s)
 CC it encodes could be used as immunoassay reagents and vaccines and to
 CC generate antibodies useful in diagnosis and passive immunotherapy for HCV
 CC infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
 CC field.) (Updated on 25-MAR-2003 to correct PI field.)
 XX
 SQ Sequence 7310 BP; 1491 A; 2217 C; 2058 G; 1540 T; 0 U; 4 Other;
 Alignment Scores:
 Pred. No.: 1.42e-241 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-930-591-2 (1-686) x AAN92106 (1-7310)
 Qy 1 MetaProteinThraLAtyAlaGlnThrArgGlyLeuGlyCysIlelleThr 20
 Db 1728 CTGGCGCCCATACGGGGTACGCCACAGACAGAGGGGCTCTAGGGTGCAATACCC 1787
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThraLa 40
 Db 1788 AGCTACTGGCGGGGACAAACCAAGTGGAGGTGAGGTGAGATTGGTCACTGCT 1847
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyHisGlyAla 60
 Db 1848 GCCCAACCTCTCTGGCAAGCTGCATCAATGGGGTGTGCTGGACTCTTACCACGGGGC 1907

QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyThrAsnValAsp 80
 Db 1908 GGAACGAGGACCATCGCGTCACCCCAAGGGTCTCTGTCATCCAGATGTATACCAATGTAGAC 1967
 QY 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1968 CAAGACCTTGTGGGCTGGCCCGCTCCGCAAGGTAGCCGCTCATTCACACCCCTGCATTCGC 2027
 QY 101 GlySerSerAspLeuTyThrValThrArgHisAlaAspValIleProValArgArg 120
 Db 2028 GGCTCTCTCGACCTTTTACCTGTCACGAGGACGCGGATGTTCATTCCTCGCGCGGGGG 2087
 QY 121 GlyAspGlyArgGlySerLeuSerProArgProIleSerTyThrLeuGlyGlySerSer 140
 Db 2088 GGTGATAGCAGGGGAGCGCTGCTGTCGCCCGGCCCATTTCTACTTTGAAAGGGTCTCTCG 2147
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 2148 GGGGGTCCGCTGTGTGCCCCCGGGGACGCGCGGGGCATATTTAGCGCGCGGGTGTGC 2207
 QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 2208 ACCCGTGGAGTGGCTAAGCGGTGACCTTATCCCTGTGGAGAACCTAGACACAACCATG 2267
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyGluVal 200
 Db 2268 AGGTCCCCGGGTTCACGGATAACTCTCTCCACAGTAGTGGCCCGAGGCTTCCAGGTG 2327
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyAla 220
 Db 2328 GCTCACCTCCATGCTCCACAGGCGGCAAGAGCAACCAAGGTCCTGGGTGCTATATGCA 2387
 QY 221 AlaGlnGlyTyThrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 2388 GCTCAGGGCTATAAGGTCTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGGTCTGGT 2447
 QY 241 AlaTyMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 2448 GCTTACATGTCACAGGCTCATGGATCGATCTTACATCAGGACCGGGGTGAGACAAAT 2507
 QY 261 ThrThrGlySerProIleThrTySerThrTyThrGlyLysPheLeuAlaAspGlyGlyCys 280
 Db 2508 ACCACTGGCAGCCCATCAGTACTCCACCTACGCAAGTTCCTTGGCCGCGGGGTGC 2567
 QY 281 SerGlyGlyAlaTyArgPheIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 2568 TCGGGGGGCGCTTAGCATTAATTTGTGACAGTGCCTCCAGGATGCCACATCC 2627
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2628 ATCTTGGGCATCGGCACCTGCTCTTGACCAAGCAGAGACTGGCGGGGCGAGACTGTTGTG 2687
 QY 321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal 340
 Db 2688 CTCGCCACGCCACCCCTCGGGTCTCGTCACTGTGCCCATCCCAACATCGAGAGGTT 2747
 QY 341 AlaLeuSerThrThrGlyGluIleProPheTyThrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2748 GCTCTGTCCACCACCGGAGAGATCCCTTTTAAAGCAAGGCTATCCCTCGAAGTAATC 2807
 QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 2808 AAGGGGGGAGACATCTCTCTGTCAATTCAAAGAGAGTCCGACGACACTCGCGCA 2867
 QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyTrArgGlyLeuAspValSerVal 400
 Db 2868 AAGTGTCTCGATGGGCATCAATGCGGTGCTTACTACCGCGTCTTGTGCTGCTCGTC 2927
 QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2928 ATCCGACACCGCGATGTTGTCTGTGTCACCAACCGATGCCCTCATGACCGGTATACC 2987
 QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

```

Db 2988 GGGGACTTTCAGCTCGGTGATAGACTGCAATACCTGTGTACCCAGACAGTCGATTCACG 3047
Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTACCATTGAGACAATCAGCTCCCCAGGATGCTGTCTCCGCACT 3107
Qy 461 GlnArgArgGlyArgThrGlyArgGlyArgProGlyIleThrArgPheValAlaProGly 480
Db 3108 CAACGTCGGGGCAGGACTGCGACGGGGGAGCCAGGCATCTACAGATTGTGGCACCGGGG 3167
Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGGCCCCCTCCGGCATGTTGCACTCGTCCGTCCTCTGTGAGTGCTATGACGAGGACTGT 3227
Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 3228 GCTTGGTATAGCTCAGCCCGCGAGACTACAGTATAGGTACGAGCGTACATGAACACC 3287
Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCGCTGTGCGAGGACCATCTTGAATTTTGGAGGGCGTCTTTACAGGCGCTC 3347
Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrIleGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCACTTCTTATCCAGACAAAGCAGAGTGGGAGAACCTTCCCTTAC 3407
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTyrAsp 580
Db 3408 CTGCTAGCGTACCAAGCCCGTGTGCGTAGGGCTCAAGCCCTCCCGCATGTTGGGAC 3467
Qy 581 GlnMetTyrPysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAAGTGTGATTGCTCCCAAGCCCAAGCCCTCCATGGGCCAACACCCCTGCTA 3527
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGCGCTGTTCAGATGAATCAACCTGACGACCCAGTCAACCAATACATC 3587
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 3588 ATGACATGCATGTGCGCGACCTGGAGTGTCTACGAGCACCTGGGTGCTGTGTCGCGGC 3647
Qy 641 ValLeuAlaAlaLeuAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTGCTGTGTCATAGTGGCAGG 3707
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCGCTGTGTCGGGAGCGGCAATCATACCTGACAGGGAAGTCTCTTACCAGAGATTG 3767
Qy 681 AspGluMetGluGluCys 696
Db 3768 GATGAGATGGAAGAGTGC 3785

RESULT 5
AAN90336
ID AAN90336 standard; DNA; 7310 BP.
XX
AC AAN90336;
XX
DT 25-MAR-2003 (revised)
DT 19-JUL-2001 (revised)
DT 01-NOV-1989 (first entry)
XX
XX Composite hepatitis C virus (HCV) cDNA.
XX Hepatitis C virus; cDNA; clone 15e; clone k9-1; probe; vaccine; ds.
XX Pan troglodytes.
XX
XX GS2212511-A.
XX

```

```

PD 26-JUL-1989.
XX
PF 18-NOV-1988; 88GB-00027024.
XX
PR 18-NOV-1987; 87US-00122714.
PR 30-DEC-1987; 87US-00139886.
PR 26-FEB-1988; 88US-00161072.
PR 26-OCT-1988; 88US-00263584.
XX
PA (CHIR ) CHIRON CORP.
XX
XX Houghton M, Choo QL, Kuo G;
XX
XX WPI: 1989-215054/30.
XX
XX P-PSDB; AAP90288.
XX
PT Hepatitis C virus gene - used for prodn. of polynucleotide probes
PT polypeptide(s) and antibodies for diagnosis, prevention and treatment of
PT infection.
XX
XX Disclosure; Fig 47; 30pp; English.
XX
XX The sequence shows a composite hepatitis C virus (HCV) cDNA, derived by
XX aligning clones K9-1 through 15e in 5'-3' direction. The cDNA encodes
XX antigens which react with antibodies in patients with non-A non-B
XX hepatitis (NANBH). The cDNA can be used to design probes, or to
XX synthesize polypeptides, which are used to diagnose HCV-induced NANBH, to
XX raise antibodies for immunoassay or treatment, or to produce vaccines.
XX See also AAP90288, and AAN90303-35. (N.B. This record was resubmitted to
XX correct errors in the sequence.) (Updated on 25-MAR-2003 to correct PR
XX field.)
XX
SQ Sequence 7310 BP; 1495 A; 2218 C; 2058 G; 1539 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,42e-241 Length: 7310
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 1 Gaps: 0

US-09-930-591-2 (1-686) x AAN90336 (1-7310)
Qy 1 MetalProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db 1728 CTGGGCGCCATCACGCGCTACGCCCCAGCAGACAGAGGGGCTCTTAGGTGCAATATCACC 1787
Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db 1788 AGCCTAATCTGGCGCGGACAAACCAAGTGGAGGTGAGGTCAGATTGTGTCACTGCT 1847
Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTyrThrValTyrHisGlyAla 60
Db 1848 GCCCAACCTTCTCTGGCAACGTGCATCAATGGGTGTGCTGGACTGTCTACACGGGGCC 1907
Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db 1908 GGACCGAGGACCATCGCTGCTACCCCAAGGGTCTGTGTCACAGATGTATACCAATGTAGAC 1967
Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db 1968 CAAAGACCTTGTGGGTGGCGCTCCGCAAGTAGCCGCTCATTGACACCCCTGCACCTTGC 2027
Qy 101 GlySerSerAspLeuTyrIleuValThrArgHisAlaAspValIleProValArgArg 120
Db 2028 GGCTCTCGGACCTTACCTGCTCAGAGGACCGCATGTCAITCCCGTCGCGCGGGCGG 2087
Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 2088 GGTGATAGAGGGGAGGCTGCTGTGCGCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG 2147
Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160

```

```

Db 2148 GGGGGTCCGCTGTTGTGCCCCCGGGGACGCGGTGGGCATATTTAGGGCGGGGTGTGC 2207
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 2208 ACCCGTGGAGTGGCTAAGCGCGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 2267
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 2268 AGGTCCCGGTGTTACGGATAACTCTCTCCACAGTAGTGGCCCCAGAGCTTCCAGGTG 2327
QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 2328 GCTCACCTCCATGCTCCACAGGAGCGGCAAAAGACCAAGGTCCCGGCTGCATATGCA 2387
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 2388 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGGCTTTGGT 2447
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaGThrIle 260
Db 2448 GCTTACATGTCNAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 2507
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 2508 ACCACTGGCAGCCCCATCACGTACTCCACCCTACCGCAAGTTCCTTGGCCGACGGGGTGC 2567
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 2568 TCGGGGCGCGCTTATGACATAATAATTTGTGACGAGTGCCACTCCACGGATGCCACATCC 2627
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 2628 ATCTGGGCATCGGCACCTGCTCTTGACCAACAGACAGACTCGGGGGCGAGACTGGTTGTG 2687
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 2688 CTCGCCACCGCACCCCTCCGGGCTCGTCACTGTGCCCCATCCACATCGAGGAGTT 2747
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 2748 GCTCTGTCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCGCCCTCGAAGTAATC 2807
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 2808 AAGGGGGGAGACATCACTCTTGTCATCAAGAGAAAGTCCGACGAACCTCGCCGCA 2867
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 2868 AAGCTGGTCCGATTTGGGCATCAATGCCGTGGCCTTACTACCGCGTCTTGACGTGTCGTC 2927
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 2928 ATCCCGACACGCGCGATGTTGTCGTGGCAACCGATGCCCTCAGACCGGCTATACC 2987
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 2988 GCGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTCAACCAGACAGTCGATTTCAGC 3047
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 3048 CTTGACCCCTACCTTCCACATTGACAAATCACGCTCCCGCAGGATGCTGCTCCCGCACT 3107
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 3108 CAACGTGGGGCAGGACTGGCAGGGGAGCCAGGCATCTACAGATTTGTGGCACCGGGG 3167
QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 3168 GAGCGCCCTCCCGCATGTTCCGACTCGTCGCTCTCTGTGAGTGCTATGACGAGGCTGT 3227
QY 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuAlaTyrMetAsnThr 520

```

```

Db 3228 GCTTGCTATAGCTCACGCCCCCGGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 3287
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 3288 CCGGGGCTTCCCGTGTGCAGGACCATCTTGAATTTGGAGGGGCTTTTACAGGCCCTC 3347
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 3348 ACTCATATAGATGCCACTTTCTATCCAGACAAAGACAGAGTGGGGAGAACCTTCCCTAC 3407
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
Db 3408 CTGGTAGCGTACCAAGCACCCTGTGCGTAGGCTCAAGCCCTTCCCCCATCGTGGGAC 3467
QY 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 3468 CAGATGTGGAGTGTTCATTTCGCTCAAGCCACCCCTCCATGGGCAACACCCCTGCTA 3527
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 3528 TACAGACTGGGCGCTGTTTCAGAAATCAATCACCCTGACGACCCAGTCACCAATAACATC 3587
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db 3588 ATGACATGCATGTGCGCCGACCTCGAGGTCTGCACGACACTGGGTCTGTTGGCGGC 3647
QY 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 3648 GTCTCGGCTGCTTTGGCGCGCTATTGCTGCTCAACAGGCTGCTGTCATAGTGGCAGG 3707
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 3708 GTCTGCTGTCTCGGAAGCCGCAATCATCTGACAGGAAGTCCCTCTACCGAGATTC 3767
QY 681 AspGluMetGluGluCys 686
Db 3768 GATGAGATGGAAGAGTGC 3785

```

RESULT 6
AAQ98221
ID AAQ98221 standard; cDNA to mRNA; 7310 BP.
XX
AC AAQ98221;
XX
DT 25-MAR-2003 (revised)
DT 15-AUG-1996 (first entry)
XX
DE Hepatitis C virus clone genome.
XX
KW Hepatitis C virus; HCV; antigen; detection; diagnosis; vaccine;
KW antibodies; immunoprophylaxis; sera; serum; ds.
XX
OS Hepatitis C virus.
XX
FN US5443965-A.
XX
PD 22-AUG-1995.
XX
PF 05-APR-1991; 91US-00681703.
XX
PR 06-APR-1990; 90US-00505611.
PR 09-OCT-1990; 90US-00594854.
XX
PA (GENE-) GENELABS INC.
XX
PI Moeckli R, Reyes GR, Kim JP;
XX
DR WPI; 1995-302120/39.
XX
PT New nucleic acids encoding hepatitis C virus antigens - used to develop
PT prods. for detection of HCV-infected sera and prodn. of vaccines and anti-
PT HCV antibodies.

PS Example 4; Fig 11; 71pp; English.

XX Hepatitis C virus (HCV) antigens can be used for detecting HCV infected
 CC sera and individuals infected with HCV. They can also be used in an anti-
 CC HCV vaccine or for the production of anti-HCV antibodies which can be
 CC used for passive immunoprophylaxis. The antigens consistently identify
 CC more HCV positive serum samples with a high degree of specificity. See
 CC AAQ98202-14 and AAH81939-51. (Updated on 25-MAR-2003 to correct PF
 CC field.) (Updated on 25-MAR-2003 to correct PR field.)

XX SQ Sequence 7310 BP; 1494 A; 2217 C; 2060 G; 1539 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,42e-241 Length: 7310
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 2 Gaps: 0

US-09-930-591-2 (1-686) x AAQ98221 (1-7310)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnInThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 1728 CTGGCGCCCATCAGCGGTACGCCAGCAGCAGCAAGGGGCTCTAGGGTGCATATATCACC 1787
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 1788 AGCTTAACCTGGCGGGGACAAACCAAGTAGGGGGTGGAGTCCAGATTGGTCCACTGCT 1847
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 1848 GCCCAACCTCTCTGGCAACGTGCATCAATGGGGTGTGTGGACTGTCTACCAAGGGGCC 1907
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 1908 GGAACGAGGACCATCGCGTCACCAAGGGGTCGTGTATCCAGATGATATACCAATGTAGAC 1967
 Qy 81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 1968 CAAGACCTTGTGGCTGGCGCGCTCGCAAGTAGTGGCGCTCATTGACACCTGCCTGTC 2027
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
 Db 2028 GGCTCCTCGGACCTTACCTGGTGCACGAGGACGCGCATGTCTCCCGTGGCGCGCGG 2087
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
 Db 2088 GGTGATAGCAGGGGACCTCTGTGTCGCCCGGCCCATTTCTTCTTGAAGGCTCTCTCG 2147
 Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 2148 GGGGTCCTGTTGTGCCCGGGGACGCGGTGGGCATATTTAGGGCGCGGTGTGC 2207
 Qy 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 2208 ACCCGTGGAGTGGTAAAGGCGGTGGACTTATCCCTGTGGAGAACCTTAGAGACCAACATG 2267
 Qy 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 Db 2268 AGTCCCCGGGTTCACGGATAACTCTCTCCACAGTAGTGGCCCGCAGAGCTTCCAGGTG 2327
 Qy 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 Db 2328 GCTCACCTCCATGCTCCCGCAGGCGGCAAAAGCACCAGGTCCCGGCTGCATATGCA 2387
 Qy 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 2388 GCTCAGGGCTATAAGTGTGTAGTACTCAACCCCTCTGTGTCTGCTGCAACACCTGGGCTTGGT 2447
 Qy 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAla-gThrIle 260
 Db 2448 GCTTACATGTCCAGGGCTCATGGATCGATCTTAACATCAGACCGGGGTGAGAACAAAT 2507

Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyCys 280
 Db 2508 ACCACTGGCAGCCCATCAGCTACTCCACCTACGCAAGTTCCTTGGCGAGGGGGTGC 2567
 Qy 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 2568 TCGGGGGCGGTATGACATATATTTGTGACGAGTGCCACTCCACGGATGCCACATCC 2627
 Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 2628 ATCTTGGGCATCGGCATCTCTTACCAAGCAGAGACTGGCGGGGAGAGACTGGTGTG 2687
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 2688 CTCGCCACCGCACCCCTCCGGCTCCGTCACTGTGCCCCATCCACATCCAGAGAGTT 2747
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
 Db 2748 GCTCTGTCCACACCGGAGAGATCCCTTTTACGCAAGGCTATCCCTCCGAGTATC 2807
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 2808 AAGGGGGGAGACATCTCTCTGTCTATCAAGAGAGAGTGGCAGAACTCCGCCA 2867
 Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 2868 AAGCTGTTCGATTTGGGCATCAATGCGTGGCTACTACCGCGTCTTTGACGTGCTC 2927
 Qy 401 IleProThrSerGlyAspValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 2928 ATCCGACACGCGGCGATGTTGTCTGTGGCAACCGATGCCCTCATACCGGTATACC 2987
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 2988 GCGGACTTCGACTCGGTGATAGACTGCAATAGGTGTGCACCCAGACAGTCAATTTCAG 3047
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 3048 CTGTACCTTACCTTCCATTGAGCAATCACTGCTCCCGCAGGATGCTCTCCCGCAT 3107
 Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 3108 CAACCTCGGGGAGAGCTGGCAGGGGAGGAGCAGCACTACAGATTTGTGGCAGCGGG 3167
 Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 3168 GAGCCCGCTCCGGCATGTTGACTCGTCCGCTCTGTGTGAGTGTATGACGAGGCTGT 3227
 Qy 501 AlaThrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 3228 GCTTGTGTAGCTCACGCCCGCAGACTACAGTACGCTACGAGCGGTACATGAACACC 3287
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 3288 CCGGGGCTTCCGCTGTGCCAGGACCATCTTGAATTTGGGAGGGCGTCTTTACAGCCTC 3347
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 3348 ACTCATATAGATGCCCATTTCTATCCAGACAAAGCAGAGTGGGGAGAACCTTCCATTAC 3407
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 3408 CTGGTAGGTACCAAGCCACCGTGTGCGCTAGGGCTCAAGCCCTCCCTCCCATCTGGGAC 3467
 Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 3468 CAGATGTGGAAGTGTGATTGCGCTCAAGCCACCTCCATGGGCGCAACCCCTGCTGA 3527
 Qy 601 TyrArgLeuGlyValAlaGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 3528 TACAGACTGGGCGCTGTTTCAGAAATGAAATGCACTGACGCCACCCAGTCAACCAATATC 3587

QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 DB 3588 ATGATCATGTCGGCCGACCTGGAGTCTCAGGAGCACCCTGGTCTCGTTGGCGGC 3647
 QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 DB 3648 GTCTCTGGTGTCTTTGGCCGCGTATTGCTGTCAACAGGCTCGTGGTCATAGTGGCGAGG 3707
 QY 661 IleValLeuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680
 DB 3708 GTCTGTCTTCTCGGAACCGCGAATCATCTCAGAGGAGTCTCTACCGAGAGTTC 3767
 QY 681 AspGluMetGluGluCys 686
 DB 3768 GATGAGATGGAAGAGTGC 3785
 RESULT 7
 AAA75296
 ID AAA75296 standard; cDNA; 8316 BP.
 AC AAA75296;
 DT 15-JAN-2001 (first entry)
 DE cDNA sequence compiled Hepatitis C virus cDNA clones.
 KW Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;
 OS Hepatitis C virus.
 FH Key Location/Qualifiers
 FT 1..8316
 FT /*tag= a
 FT /note= "partial sequence; no termination codon given"
 XX
 PN EP1034785-A2.
 XX
 PD 13-SEP-2000.
 XX
 PF 16-MAR-1990; 2000EP-00109602.
 XX
 PR 17-MAR-1989; 89US-00325338.
 PR 20-APR-1989; 89US-00341334.
 PR 18-MAY-1989; 89US-00355002.
 PR 16-MAR-1990; 90EP-00302866.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Houghton M, Choo Q, Kuo G;
 XX
 DR WPI; 2000-566891/53.
 DR P-FSDB; AAB18540.
 XX
 PT Novel composition comprising a hepatitis C virus antisense polynucleotide
 PT which is complementary to or corresponds to a sense strand of the virus
 PT genome, and selectively hybridizes to it.
 XX
 PS Example; Fig 16; 75pp; English.
 XX
 CC The specification describes a pharmaceutical composition which comprises
 CC a hepatitis C virus (HCV) antisense polynucleotide. The HCV is
 CC characterized by a positive stranded RNA genome which has 40% homology at
 CC the polypeptide level to a HCV polyprotein. The antisense polynucleotide
 CC binds to cellular polynucleotides which enhance and/or are required for
 CC viral infectivity, replicative ability or chronicity. The antisense
 CC polynucleotides may also be designed to bind with high specificity, to be
 CC of increased stability, to be stable and to have low toxicity. The
 CC composition also comprises an agent which causes viral RNA to be
 CC inactive. The composition is used for preventing HCV replication in a
 CC system. The present sequence represents a novel HCV cDNA sequence, which
 CC is used in the course of the invention
 XX

SQ Sequence 8316 BP; 1671 A; 2529 C; 2345 G; 1771 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.66e-241 Length: 8316
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: 3 Gaps: 0
 US-09-930-591-2 (1-686) x AAA75296 (1-8316)
 QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 DB 2734 CTGGCGCCCATCAGCGCTAGCCCGCAGCAGCAAGGGCCCTCTAGGTGTCATATCACC 2793
 QY 21 SerLeuThrGlyArgAspIleAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 DB 2794 AGCCTAACTGGCCGGGCAAAAACCAAGTGAGGTGAGGTCCAGATGTGTCAACTGCT 2853
 QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTTPThrValTyrHisGlyAla 60
 DB 2854 GCCCAAACTTCTGGCAACGTCATCAATGGGGTGTCTGACTGTCTACCACGGGGCC 2913
 QY 61 GlyThrArgThrIleAlaSerProIleGlyProValIleGlnMetTyrThrAsnValAsp 80
 DB 2914 GGAACGAGGACCATCGCTCACCAAGGCTCTGTCTATCCAGATGTATACCAATGTAGAC 2973
 QY 81 GlnAspLeuValGlyTTPProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 DB 2974 CAAAGACCTTGTGGCTGGCCCGCTCCGCAAGTGTAGCGCTCATTTGACACCTGCTGC 3033
 QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 DB 3034 GGTCTCTCGGACCTTTACTGTGTACGAGGACCGCATGTCTATCCCGTGGCGCGCGG 3093
 QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuGlyGlySerSer 140
 DB 3094 GGTGATAGCAGGGCGGCTGTGTGCGCCCGCGCCCATTTCTTACTTGAAGGCTTCTCG 3153
 QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 DB 3154 GGGGGTCCGCTGTTGTGCGCGCGGCGCACCGCTGGGCATATTTAGGGCGCGGTGTGC 3213
 QY 161 ThrArgGlyValAlaAlaLysAlaValAspPheIleProValGluSerLeuGluThrMet 180
 DB 3214 ACCCGTGGAGTGGCTAAGCGGTGACTTTATCCTGTGGAGAACCTTAGACACCAACATG 3273
 QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
 DB 3274 AGGTCCCGGCTGTTCACGGATAACTCTCTCCACAGTAGTGTCCCGCAGAGCTTCAGGTG 3333
 QY 201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
 DB 3334 GGTACCTCCATGCTCCACAGGCGGCAAAAGCACCAGGTCCTCGGCTGCATATGCA 3393
 QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 DB 3394 GCTCAGGCTATAGGTGTAGTACTCAACCCCTCTGTGTGTGCAACACTGGGGCTTGGT 3453
 QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 DB 3454 GCTTACATGTCCAAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAAAT 3513
 QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
 DB 3514 ACCACTGGCAGCCCATCAGTACTTCCACTACGGCAAGTCTCTTGGCAGCGCGGGTGC 3573
 QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 DB 3574 TCGGGGGCGCTTATGACATAATAATTTGTGACAGGTGCCACTCCACGGATGCCACATCC 3633
 QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320

Db 3634 ATCTTGGGCAATCGGCACTGCTTGGACCAAGAGAGACTCGGGGGGAGACTGGTGTG 3693
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 3694 CTGCCACCCGCCACCTCCGGGCTCGCTCACTGTGCCCCATCCCAACATCGAGGAGGTT 3753
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyValIleProLeuGluAlaIle 360
 Db 3754 GCTCTGTCCACCAACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 3813
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
 Db 3814 AAGGGGGGAGACATCT 3873
 Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 3874 AAGTGTGTGTCATTTGGGCATCAATGCGGCTGCTACTACCGCGTCTTGACGTGTGCTC 3933
 Qy 401 IleProThrSerGlyAspValValValValAlaThrThrAspAlaLeuMetThrGlyPheThr 420
 Db 3934 ATCCCGACCAAGCGGAGT 3993
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 3994 GGCACCTCGACTCGGTGATAGACTGCATACACGTGTGTGTGTGTGTGTGTGTGTGTGT 4053
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4054 CTGACCCCTACCTTCACCATGTAGACAACTACGCTCCCGCAGAGTGTGTGTGTGTGTGT 4113
 Qy 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
 Db 4114 CAACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGCATCTACAGATTTGTGGACCGGGG 4173
 Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 4174 GAGCGCCCTCCGGCATGTGTGACATCGCTCGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 4233
 Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4234 GCTTGGTATGACTCAGCCCGCCGCGAGACTACAGTTAGGTACGAGGCTACATGAACACC 4293
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 4294 CCGGGGCTCCCGTGTGCCAGGACATCTTGAATTTGGGAGGGCGTCTTTACAGGCGCTC 4353
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
 Db 4354 ACTCATATAGATGCCACTTCTATCCAGAACAAAGCAGAGTGGGGAGAACCTTCTCTTAC 4413
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProProSerTrpAsp 580
 Db 4414 CTGGTAGCGTACCAAGCCACCGCTGTGGCTAGGGCTCAAGGCCCTCCCCCATCGTGGGAC 4473
 Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 4474 CAGATGTGGAAGTGTGTGATTCGGCTCAAGCCCACTCCATGGGGCCCAACACCCCTGCTA 4533
 Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 4534 TACAGACTGGGGCTGTTCAGATGAATCACTACCTGACGACCCAGTCACCAATACATC 4593
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 4594 ATGACATCATGTGGCGGACCTGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4653
 Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 4654 GTCTGGCTGTGTGGCCGCGTATGTCTGTCAACAGGCTGCGTGTGTGTGTGTGTGTGTGT 4713
 Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680

Db 4714 GTCTGCTTGTCCGGGAAGCCGGCAATCATCTGACAGGAAGTCTCTCTACGAGAGTTC 4773
 Qy 681 AspGluMetGluGluCys 686
 Db 4774 GATGAGATGGAAGAGTGC 4791
 RESULT 8
 AA207656
 ID AA207656 standard; DNA; 9133 BP.
 XX AA207656;
 XX 20-MAR-2003 (revised)
 DT 08-NOV-1999 (first entry)
 XX Nucleotide sequence of HCV-1 ORF.
 XX Hepatitis C virus; HCV; J1; J7; HCV-1; non-A, non-B HCV; NANBH;
 KW HCV infection; vaccine; ds.
 XX Hepatitis C virus.
 PH Key Location/Qualifiers
 CDS 268..9132
 FT /tag= b
 FT /transl_except= (pos:1588..1589; aa:Leu)
 FT /note= "this codon has an apparent 1 nucleotide deletion,
 FT which alters the reading frame"
 FT /transl_except= (pos:1647..1650; aa:Pro)
 FT /note= "this codon has an apparent 1 nucleotide
 FT insertion, which alters the reading frame; this insertion
 FT is not indicated in the sequence present in the formal
 FT sequence listing of the specification"
 XX EP939128-A2.
 XX 01-SEP-1999.
 XX 17-SEP-1990; 99EP-00101746.
 XX 15-SEP-1989; 89US-00408045.
 PR 21-DEC-1989; 89US-00456142.
 PR 17-SEP-1990; 90EP-00310149.
 XX (OYAA/) OYA A.
 PA (CHIR) CHIRON CORP.
 XX Miyamura T, Saito I, Houghton M, Weiner AJ, Han J, Kolberg JA;
 PI Cha T, Irvine BD;
 XX WPI; 1999-480843/41.
 DR P-PSDB; AAY14975.
 XX New Hepatitis C Virus isolates, useful for diagnosis of hepatitis
 PT infections and development of vaccines.
 XX Disclosure; Fig 12; 132pp; English.
 PS The invention provides two new isolates of hepatitis C virus (HCV), J1
 CC and J7. These two isolates comprise nucleotide and amino acid sequences
 CC that are distinct from the HCV isolate HCV-1. The nucleotide sequences
 CC may be used to detect non-A, non-B HCV (NANBH) polynucleotides by
 CC hybridization for diagnosis of NANBH infections. They may also be used to
 CC screen blood donors, donated blood and blood products for this infection.
 CC The isolates may also be used to isolate other naturally occurring
 CC variants of the virus. The polypeptides may be used as a vaccine for
 CC administration to patients to protect against infection with NANBH. The
 CC present sequence represents the nucleotide sequence of HCV-1 ORF.
 CC (Updated on 20-MAR-2003 to correct PF field.) (Updated on 20-MAR-2003 to
 CC correct PR field.)
 XX Sequence 9133 BP; 1834 A; 2772 C; 2600 G; 1927 T; 0 U; 0 Other;

1	MetAlaProIleThrAlaTyrAlaGlnThrArgGlyLeuLeuGlyCysIleIleThr	20
...
3343	CTGGCGCCCATCAGCGGGTACGCCAGACACAAGGGGGCTCTAGAGGTGCATAATCACC	3402
...
21	SerLeuThrGlyArgAspLysAsnGlnValGluGlyValGlnIleValSerThrAla	40
3403	AGCTTAACCTGCCGGGCAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT	3462
...
41	AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla	60
3463	GCCCAAACTTCCTGGCAACGTCATCAATGGGGTGTCTGGACTGTCTACCAAGGGGCC	3522
...
61	GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp	80
3523	GGACGAGGACATCGGCTCACCCAGGGTCTGTCTCATCCAGATGTATACCAATGTAGAC	3582
...
81	GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys	100
3583	CAAGACCTTGTGGCTGGCCGCTCCGCAAGGTAGCCGCTCATTTGACACCCCTGCACCTGC	3642
...
101	GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg	120
3643	GGCTCTCGGACCTTTACCTGGTTCACAGGACGCCGATGTCATTCGCGTGGCGCGGGCG	3702
...
121	GlyAspGlyArgGlySerLeuLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer	140
3703	GGTGATACAGGGGACGCTGCTGTCGCCCGGCCCAATTTCTACTTGAAGAGCTCTCTGC	3762
...
141	GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys	160
3763	GGGGTCCGCTGTGTGTCGCCCGGGGGACAGCGCGTGGCATATTTAGGGCGCGGTTGTC	3822
...
161	ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet	180
3823	ACCCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGACAACCTAGAGCAACCATG	3882
...
181	ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal	200
3883	AGGTCCCGGTGTTCACGGATAACTCTCTCCACCATGTAGTCCCCACAGAGTCTCCAGGTG	3942
...
201	AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla	220
3943	GCTACCTCCATGCTCCACAGGACGGGCAAAAGCACCAAGGTCCCGGCTGCATATGCA	4002
...
221	AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly	240
4003	GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTGTGCAACACTCGGGGTTTGGT	4062
...
241	AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle	260
4063	GCTTACATGTCGANGGCTCATGGGATCGATCTTACATCAGGACCGGGGTGAGAACATT	4122
...
261	ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys	280
4123	ACCACTGGCAGCCCATCAGCTACTCCACTCAGGCAAGTTCTCTTGTCCGACGCGGGGTGC	4182
...
281	SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer	300
4183	TCGGGGGGCGCTATGACATAATATTGTGACGAGTGCCACTCCACGGATGCCACATCC	4242
...
301	IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal	320
4243	ATCTTGGGCATCGGCATGCTCTTGACCAAGCAGAGACTGGGGGGCGAGACTGTTGTG	4302

```

QY      681 AspGluMetGluGluCys 686
Db      5383 CATGAGATGGAAGAGTGC 5400

RESULT 9
AAQ10566
ID      AAQ10566 standard; DNA; 9185 BP.
XX      AC      AAQ10566;
XX      25-MAR-2003 (revised)
DT      29-APR-1991 (first entry)
XX      DE      Hepatitis C virus strain 1 DNA.
XX      KW      Hepatitis C virus; HCV-1; non-A, non-B hepatitis; HCV antigen;
XX      KW      viral infections; ss.
XX      OS      Hepatitis C virus.
XX      PN      EP414475-A.
XX      PD      27-FEB-1991.
XX      PF      21-AUG-1990; 90EP-00309120.
XX      PR      25-AUG-1989; 89US-00398667.
XX      PA      (CHIR ) CHIRON CORP.
XX      PI      Weiner AJ, Steimer KS;
XX      DR      WPI; 1991-059670/09.
XX      PT      Cell lines infected with hepatitis C virus - are used as source of
XX      PT      antigens for detection of HCV antibodies, for vaccines, and for screening
XX      PT      anti-viral agents.
XX      PS      Disclosure; Fig 1; 24pp; English.
XX      CC      This is a hepatitis C virus (HCV) composite cDNA sequence, deduced using
XX      CC      overlapping clones. a compsn. contg. the antigenic protein encoded by
XX      CC      this sequence is useful for detecting anti-HCV anti- bodies (Abs) and for
XX      CC      screening an agent which inhibits HCV replic- action. A cell line infected
XX      CC      with this virus can be used as a source of antigens. The antigen is
XX      CC      useful for preparing vaccines for treating viral infections. See also
XX      CC      AAQ10567. (Updated on 25-MAR-2003 to correct PA field.)
XX      SQ      Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1.88e-241      Length:      9185
Score:          3574.00      Matches:      672
Percent Similarity: 99.42%      Conservative: 10
Best Local Similarity: 97.96%      Mismatches:  4
Query Match:    98.78%      Indels:      0
DB:            2      Gaps:      0

US-09-930-591-2 (1-686) x AAQ10566 (1-9185)
QY      1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
Db      3395 CTGGCGCCCATCGCGGTAGCGCCAGACAGCAAGGGCCCTCCTAGGGTGCAATAACAC 3454
QY      21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3455 AGCCTAACTGGCCGGGCAAAAACCAAGTAGTGGAGGGTGAGTCCAGATTGTGTCAACTGCT 3514
QY      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIleThrValTyrHisGlyAla 60
Db      3515 GCCCAACCTTCTCGGCAACGTGTCATCAATGGGGTGTGTGGACTGTCTACACAGGGGCC 3574
QY      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80

```

```

Db      3575 GGAACGAGGACCATCGCGTCAACCAGGGTCTCTGTCTCCAGATGTATCCAAATGTAGAC 3634
QY      81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      3635 CAAGACCTTGTGGCTGGCGCTCCGCAAGGTAGCGGCTCAITGACACCTCGACCTGTC 3694
QY      101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
Db      3695 GGCTCTCGACCTTTACCTGGTCACGAGGCACCGCATGTCAITCCGTCGCGCGCGG 3754
QY      121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db      3755 GGTGATAGCAGGGCAGCCTCTGTGCGCCGCGCCATTCTTCTACTTGAAGAGGTCTCTCG 3814
QY      141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      3815 GGGGGTCCGCTGTGTGCCCGCGGGCACCGCGTGGGCATATTAGGGCCGCGGTGTC 3874
QY      161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db      3875 ACCCGTGGAGTGGCTAAGCGCGTGGACTTTATCCCTGTGGAGAACCTAGACAAACCATG 3934
QY      181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db      3935 AGTCCCCCGGTGTTACCGGATACTCTCTCCACAGTAGTGGCCCGCAGAGCTTCAGGTG 3994
QY      201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db      3995 GCTCACCTCCATGCTCCACAGGCAGCGGCAAAAGCAAGGTCGCGGTGCATATGCA 4054
QY      221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db      4055 GCTCAGGGCTATAAGGTGCTAGTACTCAACCCCTCTGTGTCTGCAACACTGGGCTTTGGT 4114
QY      241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
Db      4115 GCTTACATGTCCAAGGCTCATGGATCGATCCTTAACATCAGGACCGGGGTGAGAACAATT 4174
QY      261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db      4175 ACCACTGGCAGCCCCATCACGTACTCCACTACGCAAGTTCCTTGGCCAGCGGGGTGC 4234
QY      281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db      4235 TCGGGGGCGCTTATGACATAATTAATTTGTACGAGTGCCACTCCACGAGTCCACATCC 4294
QY      301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db      4295 ATCTTGGGCATCGGCACCTGCTCTTGACCAAGCAGAGACTGCGGGGGCGAGACTGCTTGTG 4354
QY      321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluGluVal 340
Db      4355 CTCGCCACCGCACCCCTCCGGGCTCCGTCACCTGTGCCCATCCCAACATCGAGAGGTT 4414
QY      341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db      4415 GCTCTGTCCACACCGGAGAGATCCTTTTACGCAAGGCTATCCCCCTCGAAGTAATC 4474
QY      361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysLysCysAspGluLeuAlaIle 380
Db      4475 AAGGGGGGAGACATCTCATCTTCTGTCAATCAAGAGAGAGTGGCAGAACTCGCCGCA 4534
QY      381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db      4535 AAGCTGTGCGATTGGGCATCAATGCGCTACTACCGGGCTTTCAGCTGTCCGTC 4594
QY      401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db      4595 ATCCGACAGCGCGGATGTTGTCTGCTGGCAACCGATGCCCTCATGACCGGTATACC 4654
QY      421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440

```

Db 4655 GCGACCTTCGACTCGGTGATAGACTGCAATACGTTGTGTACCCACAGACAGTCGATTTCAGC 4714
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4715 CTTGACCTTACCTTCACCATTTAGACAAATCACGCTCCCCAGGATGCTGTCTCCCGCACT 4774
 Qy 461 GlnArgArgGlyArgThrGlyValArgGlyLysProGlyIleThrArgPheValAlaProGly 480
 Db 4775 CAAGCTCGGGCAGAGACTGGCAGCGGGAGAGCCAGGCACTACAGATTTGTGGCACCGGG 4834
 Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuGluCysGluCysThrAspAlaGlyCys 500
 Db 4835 GAGCGCCCTCCGCGCATGTTCCGACTCGCTCGCTCTGTGAGTGTATGACGCGAGGTGT 4894
 Qy 501 AlaTrpTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4895 GCTTGGTATGAGCTCAGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4954
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeu 540
 Db 4955 CCGGGCTTCCCGTGTGCGAGGACCACTTGAATTTGGGAGGCGTCTTTACAGGCTC 5014
 Qy 541 ThrHisLeuAspAlaHisPheLeuSerGlnThrIlysglnSerGlyGluAsnLeuProTyr 560
 Db 5015 ACTCATATAGATGCCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCCTTAC 5074
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
 Db 5075 CTGTAGGTACCAAGCACCGTGTGGCTAGGCTCAAGCCCTCCCATCGTGGGAC 5134
 Qy 581 GlnVetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5135 CAGATGTGGAGTGTGTTGATTCGCTCAAGCCCAAGCCCTCCATGGGCAACACCCCTGCTA 5194
 Qy 601 TyrArgLeuGlyAlaValAlaGlnAspGluValThrLeuThrHisProValThrLysTyrIle 620
 Db 5195 TACAGATGGCGCTGTTCAGATGAATCACCCTGACGACCCAGTCACCAATACATC 5254
 Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
 Db 5255 ATGACATGTCATGCGCGCACCTGGAGTGTGTCAGGACACCTGGGTGCTGCTGGCGG 5314
 Qy 641 ValLeuAlaLeuAlaLeuAlaTyrCysLeuSerThrClyCysValValIleValGlyArg 660
 Db 5315 GTCTGTGCTGTTTGGCGCGGTATGCTGTCAACAGGCTGCGTGTGTCATAGTGGCAGG 5374
 Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 5375 GTCTGCTTGTCCGGAAGCCGCAATCATACCTGACAGGGAAGTCTCTACCGAGATTTC 5434
 Qy 681 AspGluVetGluGluCys 686
 Db 5435 GATGAGATGAAGAGTGC 5452

RESULT 10

AAA75297

ID AAA75297

AC AAA75297;

XX AAA75297;

DT 15-JAN-2001 (first entry)

XX Sense strand of HCV encoding a polyprotein.

XX Hepatitis C virus; HCV; antisense polynucleotide; polyprotein;

XX viral infectivity; viral replication; ds.

XX Hepatitis C virus.

OS

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX EP1034785-A2.
 PN 13-SEP-2000.
 XX 16-MAR-1990; 2000EP-00109602.
 XX 17-MAR-1989; 89US-00325338.
 PR 20-APR-1989; 89US-00341334.
 PR 18-MAY-1989; 89US-00355002.
 PR 16-MAR-1990; 90EP-00302866.
 XX (CHIR) CHIRON CORP.
 PA
 XX Houghton M, Choo Q, Kuo G;
 PI WPI; 2000-566891/53.
 XX P-PSDB; AAB18541.
 DR
 XX Novel composition comprising a hepatitis C virus antisense polynucleotide which is complementary to or corresponds to a sense strand of the virus genome, and selectively hybridizes to it.
 PT
 PT
 PT
 PS Example; Fig 17; 75pp; English.
 XX The specification describes a pharmaceutical composition which comprises a hepatitis C virus (HCV) antisense polynucleotide. The HCV is characterized by a positive stranded RNA genome which has 40% homology at the polypeptide level to a HCV polypeptide. The antisense polynucleotide binds to cellular polynucleotides which enhance and/or are required for viral infectivity, replicative ability or chronicity. The antisense polynucleotides may also be designed to bind with high specificity, to be of increased stability, to be stable and to have low toxicity. The composition also comprises an agent which causes viral RNA to be inactive. The composition is used for preventing HCV replication in a system. The present sequence represents a novel HCV cDNA sequence, which is used in the course of the invention

SQ Sequence 9185 BP; 1849 A; 2790 C; 2608 G; 1938 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.88e-241	Length:	9185
Score:	3574.00	Matches:	672
Percent Similarity:	99.42%	Conservative:	10
Best Local Similarity:	97.96%	Mismatches:	4
Query Match:	98.78%	Indels:	0
DB:	3	Gaps:	0

US-09-930-591-2 (1-686) x AAA75297 (1-9185)

Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3395 CTGGCGCCCATCACGGCGGTACGCCACAGACAGAGGGGCGCTCTAGGGTGATAATCACC 3454
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
 Db 3455 AGCCTACTGTCGGCGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTCTCACTGCT 3514
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 3515 GCCCAACCTTCTCTGGCAACGTGCATCAATGGGTGTCTGGACTGTCTACCGGGGCC 3574
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 3575 GGACGAGGACCATCGGTACCCCAAGGTCCTGTATCCAGATGTATACCAATGTAGAC 3634
 Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3635 CAAGACCTTGTGGGCTGCGCCGCTCCGCAAGGTAGCGGCTCATTGACACCCCTGCAC 3694
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
 Db 3695 GCCTCTCGGACCTTTACCTGGTCCAGGACGCCGATGTCTATCCCGTGGCGCGCGG 3754

Key Location/Qualifiers
 CDS 320..9184
 /*tag= a
 /note= "partial sequence; no termination codon given"

```

QY 121 GluAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuLysGlySerSer 140
Db 3755 GGTGATAGCAGGCGACCTGCTGTCGCCCGCGGCCATTTCTTACITTAAGGCTCTCTCG 3814
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db 3815 GGGGGTCCGGCTGTGTGCGCGCGGGCACCGCGTGGGCATATTTAGGGCGCGGTGTC 3874
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db 3875 ACCCGTGGAGGGGTAGAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGACAACCATG 3934
QY 181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db 3935 AGGTCCCGGTGTTCCACGGATAACTCTCTCCACCACTAGTGGCCCAAGACTTCCAGGTG 3994
QY 201 AlaHisIleuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
Db 3995 GCTCACCTCCATGCTCCACAGGCGAGCGGCAAGACCAAGGTCGCCGTGCTATGCA 4054
QY 221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
Db 4055 GCTCAGGGCTATAGGTGCTAGTACTCAACCCCTCTGTTGCTGCAACACTGGGCTTTGGT 4114
QY 241 AlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValAlaArgThrIle 260
Db 4115 GCTTACATGTCCAAAGGCTCATGGATCGATCCTAACATCAGGACCGGGGTGAGAACAAIT 4174
QY 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
Db 4175 ACCACTGGACGCCCATCAGTACTCCACCTACGGCAAGTTCCTTGGCCGACGGGGTGC 4234
QY 281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
Db 4235 TCGGGGGCGGCTTATGACATAATAATTTGTGACAGTGCACCTCCACGGATGCCACATCC 4294
QY 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
Db 4295 ATCTTGGGCATCGGCATGCTCTTGACCAAGACGAGACTCGGGGGCGAGACTGGTGTG 4354
QY 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
Db 4355 CTGCCACCGCCACCTCCGGGCTCGTCACTGTGCCCCATCCCAACATCGAGGAGT 4414
QY 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
Db 4415 GCTCTGCCACCCACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCCCTCGAAGTAATC 4474
QY 361 LysGlyGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
Db 4475 AAGGGGGGAGACATCTCATCTTCTGTCATTCAAAGAGAAGTGGCGACGACTCGCGCA 4534
QY 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
Db 4535 AAGTGTGTCGATTTGGCATCAATCGCGTGGCTTACTACCGCGTCTTTGACGTCCGTC 4594
QY 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
Db 4595 ATCCCGACCAAGCGGAGATGTGCTGCTGCGAACCATGCTCCCTCATGACCGGCTATACC 4654
QY 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
Db 4655 GGGGACTTCGACTCGGTGATAGACTGCAATACGTGTGTACCCAGACAGTCGATTTCCAGC 4714
QY 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
Db 4715 CTTGACCGCTACCTTACCATTGAGACAAATCAGCTCCCGCAGGATGCTGTCTCCCGCACT 4774
QY 461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
Db 4775 CAACGTGGGGCAGGACTGGCAGGGGGAAGCCAGGCACTTACAGATTTGTGGCACCGGG 4834

```

```

QY 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
Db 4835 GAGCGCCCTCCGGCATGTTTCGACTCGTCCTGAGTGTGCTATGACGCGGCTGT 4894
QY 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
Db 4895 GCTTGGTATGAGCTCACGCCCGCCGAGACTACAGTTAGGCTACGAGCGTACATGAACACC 4954
QY 521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
Db 4955 CCGGGCTTCCCGTGTGCCAGGACCATCTTGAATTTGGAGGGGCTCTTTACAGGCTTC 5014
QY 541 ThrHisIleAspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyr 560
Db 5015 ACTCATATAGATGCCACTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTCTTAC 5074
QY 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db 5075 CTGGTAGCGTACCAAGCACCGGTGTGCGTAGGGCTCAAGCCCTCCCCCATCGTGGGAC 5134
QY 581 GlnMetTyrLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db 5135 CAGATGTGGAAGTGTTCGCTCAAGCCCACTCCATGGGCCCAACACCTTCTGCTA 5194
QY 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db 5195 TACAGACTGGGCGCTGTTCAAGATGAATCAACCTTCAGCGCACCCAGTACCCAAATACATC 5254
QY 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTyrValLeuValGlyGly 640
Db 5255 ATGACATGCATGTCGGCGGACCTGGAGGTGCTCAGCAGCAGCCTGGGTGCTCGTGGCGGC 5314
QY 641 ValLeuAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db 5315 GTCTGGTGTGCTTGGCGCGCTATTGCTGTCAACAGGCTGGGTGCTCATAGTGGCGAG 5374
QY 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db 5375 GTCTGCTTGTCCGGGAAGCCGCAATCATCTGACGGGAAGTCTCTTACCGAGGTTCT 5434
QY 681 AspGluMetGluGluCys 686
Db 5435 GATGAGATGGAAGAGTGC 5452

```

RESULT 11

AAT12710
ID 'AAT12710 standard; cDNA; 9401 BP.
XX
AC AAT12710;
XX
DT 25-MAR-2003 (revised)
DT 15-MAY-1996 (first entry)
XX
DE Hepatitis C virus polyprotein.
XX
KW Non-A non-B hepatitis virus; NANBHV; HCV; antigen; detection; diagnosis;
KW antibodies; ds.
XX
OS Hepatitis C virus.
FH Key Location/Qualifiers
FT CDS 342..9378
FT /*tag= a
XX
PN EP693687-A1.
XX
PD 24-JAN-1996.
XX
PF 03-APR-1991; 95EP-00114016.
XX
PR 04-APR-1990; 90US-00504352.
PA (CHIR) CHIRON CORP.


```

Db      5097 CTGGTAGGTTACCAAGCCAGCGTGGCTAGGCTCAAGCCCTCCCCATCGTGGAC 5156
QY      581 GlnMetTrpLysCysLeuLleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db      5157 CAGATGTGGAAGTGTTCATTTCGCTCAAGCCACCCCTCCATGGCCCAACACCCCTGCTA 5216
QY      601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db      5217 TACAGCTGGGGCGTGTTCAGATGAATCAATCAATCAATCAATCAATCAATCAATCAATC 5276
QY      621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGlyGly 640
Db      5277 ATGACATCATGTCCGCCGACCTGGAGTCTGTCACGACCTGGTGTCTGTGGCGGC 5336
QY      641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db      5337 GTCTTGGCTGTTCGGCGGTATTGCTGTCAACAGGCTGCGTGGTCATAGTGGCGAG 5396
QY      661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db      5397 GTGCTGTTCGGGAGCCGCAATCAATCAATCAATCAATCAATCAATCAATCAATC 5456
QY      681 AspGluMetGluGluCys 686
Db      5457 GATGAGATGGAAGAGTGC 5474

RESULT 12
AAT99981
ID   AAT99981 standard; DNA; 9401 BP.
XX
AC   AAT99981;
XX
DT   25-MAR-2003 (revised)
DT   16-MAR-1998 (first entry)
DE   HCV polypeptide coding sequence.
XX
KW   PCR primer; amplify; HCV; hepatitis c virus; antigen combination; NS3;
KW   C domain; S domain; NS; HCV polypeptide; anti-HCV antibody; detection;
KW   NS4; ds.
XX
OS   Hepatitis C virus.
XX
PH   Key Location/Qualifiers
FT   CDS 342..9377
FT   FT /*tag= a
XX
PN   US5683864-A.
XX
PD   04-NOV-1997.
XX
PF   07-JUL-1992; 92US-00910760.
XX
PR   18-NOV-1987; 87US-00122714.
PR   30-DEC-1987; 87US-00139886.
PR   26-FEB-1988; 88US-00161072.
PR   06-MAY-1988; 88US-00191263.
PR   26-OCT-1988; 88US-00263584.
PR   14-NOV-1988; 88US-00271450.
PR   17-MAR-1989; 89US-00325338.
PR   20-APR-1989; 89US-00341334.
PR   21-APR-1989; 89US-00353896.
PR   18-MAY-1989; 89US-00355002.
PR   04-APR-1990; 90US-00504352.
XX
PA   (CHIR ) CHIRON CORP.
XX
XX   Kuo G, Houghton M, Choo Q;
XX
DR   WPI; 1997-548976/50.
DR   P-PSDB; AAW34480.
XX
XX

```

```

PT      Combination of three hepatitis C virus antigens - used for detection of
XX      specific antibodies to diagnose infection.
PS      Disclosure; Col 25-46; 57pp; English.
XX
CC      This sequence represents the Hepatitis C virus polypeptide coding
CC      sequence. Fragments of this sequence can be amplified and used in the
CC      combination of HCV antigens of the invention. The HCV antigen combination
CC      comprises an antigen (Ag1) comprising the C domain (i.e. amino acids (aa)
CC      1-120 of the HCV polypeptide), or its immunologically reactive fragment
CC      containing at least 8 aa. It also comprises two additional antigens from
CC      two different polypeptide domains, including at least 8 aa from the NS3,
CC      NS4, S or NS5 domains of the polypeptide, corresponding, respectively, to
CC      aa 1050-1640; 1640-2000; 120-400 and 2000-3011 of the HCV polypeptide.
CC      Alternatively, Ag1 contains at least 8 aa from the 1-122 or 9-177 aa
CC      regions of the HCV polypeptide. These antigen combinations are used
CC      diagnostically to detect anti-HCV antibodies, using any standard
CC      immunoassay format. These antigen combinations have a broader range of
CC      reactivity with antibodies than any antigen individually. (Updated on 25-
CC      MAR-2003 to correct PR field.)
XX
SQ      Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 1 93e-241 Length: 9401
Score: 3574.00 Matches: 672
Percent Similarity: 99.42% Conservative: 10
Best Local Similarity: 97.96% Mismatches: 4
Query Match: 98.78% Indels: 0
DB: 2 Gaps: 0

```

US-09-930-591-2 (1-686) x AAT99981 (1-9401)

```

QY      1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleLeuThr 20
Db      3417 CTGGCGCCCATCATCGCGCTAGCCGCCAGCAGCAGAGGGGCTCTAGGTGCTAATATCACC 3476
QY      21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
Db      3477 AGCTTAATCTGGCGGGGACAAACCAAGTGGAGGTGAGGTCCAGATTGTGCACTGCT 3536
QY      41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
Db      3537 GCCCAAACTCTCTGGCAACGTGCATCAATGGGTGTGTGGACTGTCTACCAAGGGGCC 3596
QY      61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
Db      3597 GGAACGAGACCAATCGCTCACCCCAAGGTCTCTGTCATCCAGATGTATACCAATGTAGAC 3656
QY      81 GlnAspLeuValGlyTyrProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
Db      3657 CAAGACCTTGTGGGTGGCGCTCCGCAAGGTAGCGCTCATTCACACCTTGCACTTGC 3716
QY      101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120
Db      3717 GGCCTCTCGACCTTACCTGCTCACGAGGACGCGCATGTCAATCCGTCGCGCGGGCGG 3776
QY      121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db      3777 GGTGATAGCAGGGGAGCTCTGTGCGCCCGGCCCATTTCTTACTTGAAGGCTCTCTCG 3836
QY      141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
Db      3837 GGGGTCGCTGTGTGTCGCCCGGGGCACGCGCTGGGCATATTTAGGGCGCGGTGTGC 3896
QY      161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
Db      3897 ACCGTGGAGTGGCTAAGCGGTGGACTTTATCCCTGTGGAGAACCTTAGAGAACCAACATG 3956
QY      181 ArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnVal 200
Db      3957 AGGTCCCGGTGTTTACCGGATAACTCTCTCCACCGTAGTGCCTCCAGAGGCTTCACGGTG 4016

```


Db 5337 GTCTGCTGCTTTGGCGCGGTATGCTGTCACAGCGTGGTGTATAGTGGCGAGG 5396
 Qy 661 lIleValleuSerGlyLysProAlaIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 5397 GTCTGCTGCTTTGGCGCGGTATGCTGTCACAGCGTGGTGTATAGTGGCGAGG 5456
 Qy 681 AspGluMetGluGluCys 686
 Db 5457 GATGAGATGGAGAGTGC 5474
 RESULT 14
 AAD35043
 ID AAD35043 standard; cDNA; 9401 BP.
 XX
 AC AAD35043;
 XX
 DT 16-JUL-2002 (first entry)
 XX
 DE Hepatitis C virus (HCV) polyprotein cDNA.
 XX
 KW Hepatitis C virus; HCV; antigen; C domain; polyprotein; NS3 domain;
 KW NS4 domain; S domain; NS5 domain; gene; ss.
 XX
 OS Hepatitis C virus.
 XX
 FH Key Location/Qualifiers
 FT CDS 342..9377
 FT /tag= a
 FT /product= "HCV polyprotein"
 FT /transl_except= (pos:366..368, aa:Xaa)
 FT /note= "Xaa equals Lys or Arg"
 FT /transl_except= (pos:372..374, aa:Xaa)
 FT /note= "Xaa equals Thr or Asn"
 FT /transl_except= (pos:867..869, aa:Xaa)
 FT /note= "Xaa equals Thr or Ile"
 FT /transl_except= (pos:1341..1343, aa:Xaa)
 FT /note= "Xaa equals Val or Met"
 FT /transl_except= (pos:2148..2150, aa:Xaa)
 FT /note= "Xaa equals Ile or Leu"
 FT /transl_except= (pos:2883..2885, aa:Xaa)
 FT /note= "Xaa equals Asn or Tyr"
 FT /transl_except= (pos:3681..3683, aa:Xaa)
 FT /note= "Xaa equals Ser or Pro"
 FT /transl_except= (pos:3690..3692, aa:Xaa)
 FT /note= "Xaa equals Thr or Ser"
 FT /transl_except= (pos:4167..4169, aa:Xaa)
 FT /note= "Xaa equals Leu or Pro"
 FT /transl_except= (pos:4323..4325, aa:Xaa)
 FT /note= "Xaa equals Val or Gly"
 FT /transl_except= (pos:4701..4703, aa:Xaa)
 FT /note= "Xaa equals Tyr or Cys"
 FT /transl_except= (pos:4752..4754, aa:Xaa)
 FT /note= "Xaa equals Ser or Thr"
 FT /transl_except= (pos:5970..5972, aa:Xaa)
 FT /note= "Xaa equals Gly or Glu"
 FT /transl_except= (pos:6183..6185, aa:Xaa)
 FT /note= "Xaa equals His or Leu"
 FT /transl_except= (pos:6186..6188, aa:Xaa)
 FT /note= "Xaa equals Cys or Ser"
 FT /transl_except= (pos:6402..6404, aa:Xaa)
 FT /note= "Xaa equals Gly or Val"
 FT /transl_except= (pos:7386..7388, aa:Xaa)
 FT /note= "Xaa equals Ser or Thr"
 FT /transl_except= (pos:7494..7496, aa:Xaa)
 FT /note= "Xaa equals Phe or Tyr"
 FT /transl_except= (pos:7497..7499, aa:Xaa)
 FT /note= "Xaa equals Ala or Ser"
 FT /transl_except= (pos:7845..7847, aa:Xaa)
 FT /note= "Xaa equals Phe or Leu"
 FT /transl_except= (pos:8409..8411, aa:Xaa)
 FT /note= "Xaa equals Gly or Arg"
 FT /transl_except= (pos:9102..9104, aa:Xaa)

FT /note= "Xaa equals Gly or Arg"
 FT /transl_except= (pos:9327..9329, aa:Xaa)
 FT /note= "Xaa equals pro or Leu; These translational
 FT exceptions occur while decoding for the alternative
 FT version of HCV polyprotein (AAE22052)"
 XX
 FN US6312889-B1.
 XX
 PD 06-NOV-2001.
 XX
 PF 12-MAY-1995; 95US-00440549.
 XX
 PR 04-APR-1990; 90US-00504352.
 PR 07-JUL-1992; 92US-00910760.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Houghton M, Choo Q, Kuo G;
 XX
 DR WPI; 2002-040268/05.
 DR P-PSDB; AAE22049.
 XX
 PT Combination of hepatitis C viral (HCV) antigens, useful in improved
 PT immunoassay for detecting HCV antibodies.
 XX
 PS Example 1; Fig 1; 58pp; English.
 XX
 CC The invention relates to combination of hepatitis C viral (HCV) antigens
 CC that have a broader range of immunological reactivity than any single HCV
 CC antigen. The combinations consist of an antigen from the C domain of the
 CC HCV polyprotein, and at least one additional HCV antigen from either the
 CC NS3 domain, the NS4 domain, the S domain, or the NS5 domain and are in
 CC the form of fusion protein, a simple physical mixture, or the individual
 CC antigens commonly bound to a solid matrix. The combinations of antigens
 CC provides broad range immunoassays for anti-HCV antibodies. The invention
 CC therefore provides a method for detecting antibodies to HCV in a mammal
 CC suspected of containing such antibodies. The present sequence is HCV
 CC polyprotein cDNA
 XX
 SQ Sequence 9401 BP; 1883 A; 2860 C; 2673 G; 1985 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1.93e-241 Length: 9401
 Score: 3574.00 Matches: 672
 Percent Similarity: 99.42% Conservative: 10
 Best Local Similarity: 97.96% Mismatches: 4
 Query Match: 98.78% Indels: 0
 DB: Gaps: 0
 US-09-930-591-2 (1-686) x AAD35043 (1-9401)
 Qy 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
 Db 3417 CTGGGGCCCATCACCGCGTACGCCCCAGCAGACAGGGGCTCTAGGTGCATATCACC 3476
 Qy 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThAla 40
 Db 3477 AGCCTAACTGGCGGGACAAAACCAAGTGGAGGTGAGGTCCAGATTGTGTCAACTGCT 3536
 Qy 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHisGlyAla 60
 Db 3537 GCCCAACCTTCCTGGCACGTCATCAATGGGGTGTGCTGGACTGTCTACCACGGGCCC 3596
 Qy 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
 Db 3597 GGAACGAGGACCATCGCGTCACCCAAAGGTCTCTCATCCAGATGTATACCAATGTAGAC 3656
 Qy 81 GlnAspLeuValGlyTrpProAlaProGlnGlyAlaArgSerLeuThrProCysThrCys 100
 Db 3657 CAAGACCTTGTGGGTGGCCCGCTCGGACAGGTAGCCGCTCATTGACACCTGCACCTTGC 3716
 Qy 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArgArg 120

Db 3717 GGCTCTCGACCTTTACCTGGTCACGAGGACGCCGATGCTATTCCTCGTGCGCGCGG 3776
 Qy 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrIleuIysGlySerSer 140
 Db 3777 GGTGATAGAGGGGACGCCCTGCTGTGCGCCCGGCCCAATTTCTTACTTGAAGGCTCCTCG 3836
 Qy 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
 Db 3837 GGGGGTCCGCTGTTGTGCCCGCGGGGCACGCCGTGGGCATATTTAGGGCCGCGGTGTC 3896
 Qy 161 ThrArgGlyValAlaIysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
 Db 3897 ACCGCTGGAGTGGCTAAGCGGGTGAATTTATCCCTGTGGAGAACCCTAGACACAACCATG 3956
 Qy 181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnInsrTyrGlnVal 200
 Db 3957 AGGTCCCGGTGTTCAGGATACTCTCTCCACAGTAGTGCCACAGCTTCACAGGTG 4016
 Qy 201 AlaHisLeuHisAlaProThrGlySerGlyIysSerThrIysValProAlaIaTyrAla 220
 Db 4017 GCTCACCTCCATGCTCCACAGGCGGCGGCAAGACCAAGGTCGCGGTGCATATGCA 4076
 Qy 221 AlaGlnGlyTyrIysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
 Db 4077 GCTCAGGGCTATAAGTGTCTAGTACTCAACCCCTCTGTTCTCTCAACACTGGGCTTTGGT 4136
 Qy 241 AlaTyrMetSerIysAlaHisGlyIleAspProAsnIleArgThrGlyValArgThrIle 260
 Db 4137 GCTTACATGTCCAAGGCTCATGGATCGATCTTAACATCAGGACCGGGGTGAGAACAT 4196
 Qy 261 ThrThrGlySerProIleThrTyrSerThrTyrGlyIysPheLeuAlaaspGlyIysCys 280
 Db 4197 ACCACTGGACGCCCATCATCGTACTCCACTACCGCAAGTCTCTGCCACGCGGGTGC 4256
 Qy 281 SerGlyIysAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
 Db 4257 TCGGGGGCGCTTATGACATAATAATTTGTGACAGTGGCCACTCCACGATGCCACATCC 4316
 Qy 301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
 Db 4317 ATCTTGGGCATCGGCATGTCCTTTGACCAAGACAGACTCGGGGGCGAGACTGGTGTG 4376
 Qy 321 LeuAlaThrAlaThrProGlySerValThrValProHisProAsnIleGluGluVal 340
 Db 4377 CTGCGCACCGCCACCCCTCGGGCTCGCTCAGTGTGCCCATCCCAACATCGAGAGTT 4436
 Qy 341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyIysAlaIleProLeuGluAlaIle 360
 Db 4437 GCTCTGTCCACACCGGAGAGATCCCTTTTACGGCAAGGCTATCCCTCGAAGTAATC 4496
 Qy 361 LysGlyGlyArgHisLeuIlePheCysHisSerIysIysIysCysAspGluLeuAlaIa 380
 Db 4497 AAGGGGGGAGACATCTCATCTTCTGTCTCATTTCAAGAAGAAGTGCACCACTCGCGGCA 4556
 Qy 381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
 Db 4557 AAGCTGGTGCATTTGGGCATCAATCGGTGGCCCTACTACCGGGTCTTGACGTGTCGTC 4616
 Qy 401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
 Db 4617 ATCCCGACCGCGGCGATGTTGCTGCGTGGCAACCGATGCCCTCATCGCGCTATACC 4676
 Qy 421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
 Db 4677 GGGGACTTCGACTCGGTGATAGTGCATACGTGTGTCAACGACACATCGATTTCAGC 4736
 Qy 441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
 Db 4737 CTTGACCCCTACCTTCACTACCTAGATGCATACGTGTGTCAACGACACATCGATTTCAGC 4796
 Qy 461 GlnArgArgGlyArgThrGlyArgGlyIysProGlyIleTyrArgPheValAlaProGly 480
 Db 4797 CAACGTTCGGGGCAGGACTGGCAGGGGGGAAGCCAGGCAATCTACAGATTTGTGGCACCGGG 4856

Qy 481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
 Db 4857 GAGCGCCCTCCCGCATGTTCTGACTGCTCGCTCTGAGTGTGATAGCGAGCGCTGT 4916
 Qy 501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
 Db 4917 GCTTGATAGCTCACGCCCGCGAGACTACAGTTAGCTTACGAGCGTACATGAACACC 4976
 Qy 521 ProGlyLeuProValCysGlnAspHisLeuGluPheThrGluGlyValPheThrGlyLeu 540
 Db 4977 CCGGGGCTTCCCGTGTGCAGGACCATCTTGAATTTTGGGAGGCGCTTTTACAGGCGTC 5036
 Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIysGlnSerGlyGluAsnLeuProTyr 560
 Db 5037 ACTCATATAGATGCCCATTTCTATCCAGACAAAGCAGAGTGGGAGAACCTTCTTAC 5096
 Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerThrPasp 580
 Db 5097 CTGGTAGGTACCAAGCCACCGTGTGCGTAGGGCTCAAGGCCCTCCGCCCATCGTGGAC 5156
 Qy 581 GlnMetTyrIysCysLeuIleArgLeuIysProThrLeuHisGlyProThrProLeuLeu 600
 Db 5157 CAGATGTGGAAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCAACCCCTGCTA 5216
 Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrIysTyrIle 620
 Db 5217 TACAGACTGGGCGCTGTTTCAAGATGAATCAACCTGACGACCCAGTCACCAATACATC 5276
 Qy 621 MetThrCysMetSerAlaAspLeuGluValThrSerThrThrProValLeuValGlyGly 640
 Db 5277 ATGACATGATGTGGCGACCTGGAGTCTGACGAGCCTCGTGGTCTCGTGGCGGC 5336
 Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
 Db 5337 GTCCGTGGCTGCTTTGGCGCGTATTGCTGTCAACAGGCTCGGTGTCATAGTGGCAGG 5396
 Qy 661 IleValLeuSerGlyIysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
 Db 5397 GTCGTCTTGTCCGGAGCCGGCAATCATACCTGACAGGGAAGTCTCTTACCGAGAGTTC 5456
 Qy 681 AspGluMetGluGluCys 696
 Db 5457 GATGAGATGGAGAGTGC 5474
 RESULT 15
 ID AAN92103
 ID AAN92103 standard; DNA; 6905 BP.
 XX
 AC AAN92103;
 XX
 DT 25-MAR-2003 (revised)
 DT 02-MAR-1990 (first entry)
 XX
 DE Combined open reading frames of the hepatitis C virus (HCV) cDNAs from
 clones 12f through 15e.
 XX
 KW Hepatitis C virus; HCV; non-A, non-B hepatitis; NANBH.
 XX
 OS Hepatitis C virus.
 FH Key Location/Qualifiers
 CDS 3..6905
 FT /*tag= a
 FT
 XX EP318216-A.
 XX
 PD 31-MAY-1989.
 XX
 PF 18-NOV-1988; 88EP-00310922.
 XX
 PR 18-NOV-1987; 87US-00122714.
 PR 30-DEC-1987; 87US-00139886.

```

PR 26-FEB-1988; 88US-00161072.
PR 06-MAY-1988; 88US-00191263.
PR 26-OCT-1988; 88US-00263584.
PR 14-NOV-1988; 88US-00271450.
XX (CHIR ) CHIRON CORP.
XX (CHIR ) CHIRON CORP.
XX Houghton M, Choo QL, Kuo G;
XX WPI; 1989-159274/22.
XX P-PSDB; AAP92047.
XX Purified hepatitis C virus - and associated nucleic acids and
XX polypeptide(s).
XX Claim 3; Fig 32-1 - 32-7; 139pp; English.
XX It is a double-stranded nucleotide sequence of the open reading frame
XX (ORF) (tag a) extending through clones 12f to 15e of hepatitis C virus
XX (HCV) cDNA. It can be used to make oligomeric DNA hybridisation probes to
XX detect the presence of HCV nucleic acids in samples. The polypeptide(s)
XX it encodes could be used as immunoassay reagents and vaccines and to
XX generate antibodies useful in diagnosis and passive immunotherapy for HCV
XX infection/non-A, non-B hepatitis. (Updated on 25-MAR-2003 to correct PR
XX field.) (Updated on 25-MAR-2003 to correct PI field.)
XX SQ Sequence 6905 BP; 1421 A; 2082 C; 1346 G; 1456 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2,16e-241 Length: 6905
Score: 3571.00 Matches: 671
Percent Similarity: 99.42% Conservative: 11
Best Local Similarity: 97.81% Mismatches: 4
Query Match: 98.70% Indels: 0
DB: 1 Gaps: 0

US-09-930-591-2 (1-686) x AAN92103 (1-6905)
QY 1 MetAlaProIleThrAlaTyrAlaGlnGlnThrArgGlyLeuLeuGlyCysIleIleThr 20
DB 1203 CTGGCGCCCATCAGGGGTACGCCCGACGACAGCAAGGGGCTCTCTAGGGTGCATATCACC 1262
QY 21 SerLeuThrGlyArgAspLysAsnGlnValGluGlyGluValGlnIleValSerThrAla 40
DB 1263 AGCCTAACTGGCCGGGACAAACCAAGTGGAGGTCAGGTCCTCCAGATTGTGCAACTGCT 1322
QY 41 AlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysIleThrValTyrHisGlyAla 60
DB 1323 GCCCAAACTTCCTGGCAACGTGATCAATGGGGTGTGGAGTGTCTACCAACGGGGCC 1382
QY 61 GlyThrArgThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAsp 80
DB 1383 GGAACGAGGACCATCGGTGACCCCAAGGTCTGTCTATCCAGATGTATACCAATGTAGAC 1442
QY 81 GlnAspLeuValGlyTyrProAlaProGlnGlnAlaArgSerLeuThrProCysThrCys 100
DB 1443 CAAGACCTTGTGGGTGGCGTCCCGCAAGGTAGCGCTCAATGACACCCCTGCACCTTGC 1502
QY 101 GlySerSerAspLeuTyrLeuValThrArgHisAlaAspValIleProValArgArg 120
DB 1503 GGCTCCTCGGACCTTACCTGTGCAGGACGCCCGCATGTCTTCCTGTCGCGCGCGG 1562
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
DB 1563 GGTGATAGAGGGGAGCCTCTCTCGCCCGCGGCTATTTCTACTTGAAGGCTCTCTCG 1622
QY 141 GlyGlyProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaValCys 160
DB 1623 GGGGGTCCGCTGTGTGGCCCGGGGACGCCCGTGGGCATATTATTAGGGCGCGGTGTGC 1682
QY 161 ThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrThrMet 180
DB 1682 GGGGGTCCGCTGTGTGGCCCGGGGACGCCCGTGGGCATATTATTAGGGCGCGGTGTGC 180

1683 ACCCGTGGAGTGGCTAAGGGGTGGAGCTTTATCCCTGTGGAGAACCTAGAGACAACCATG 1742
181 ArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrGlnVal 200
1743 AGGTCCCGGTGTTACCGGATAACTCTCTCCACAGTAGTGGCCAGAGCTTCCAGGTG 1802
201 AlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProAlaAlaTyrAla 220
1803 GCTCACTTCCATGCTCCACAGCGAGGCAAGCAAGGTCCTCCGGTGTGCATATGCA 1862
221 AlaGlnGlyTyrLysValLeuValLeuAsnProSerValAlaAlaThrMetGlyPheGly 240
1863 GCTCAGGGCTATAAGGTGTAGTACTCAACCCCTCTGTGTGCAACACATGGGGTTTGGT 1922
241 AlaTyrMetSerLysAlaHisGlyLeuAspProAsnIleArgThrGlyValArgThrIle 260
1923 GCTTACATGTCGAAGGCTCATGGATGATCTTAACATCAGGACCGGGGTGAGAACATT 1982
261 ThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCys 280
1983 ACCACTGGCAGCCCATCAGCTACTCCACCTACGCAAGTTCCTTGGCGACGGCGGTGC 2042
281 SerGlyGlyAlaTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSer 300
2043 TCGGGGGCGGTATGACATATAATTTTGACACATGCCACTCCAGGATGCCACATCC 2102
301 IleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVal 320
2103 ATCTTGGSCATCGGCACCTGCTCTTACCAAGCAGAGACTGGGGGGGAGACTGGTTGTG 2162
321 LeuAlaThrAlaThrProProGlySerValThrValProHisProAsnIleGluVal 340
2163 CTCGCCACCGCACCCCTCCGGGCTCCGTCACTGTGCCCATCCCAACATCGAGAGGTT 2222
341 AlaLeuSerThrThrGlyGluIleProPheTyrGlyLysAlaIleProLeuGluAlaIle 360
2223 GCTCTGTCCACCCGAGAGATCCCTTTTACGGCAAGGCTATCCCTCCGAAGTAATC 2282
361 LysGlyArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAla 380
2283 AAGGGGGGAGACATCTCATCTCTGTCATTCAAGAAGAGAGTCGACGAACTCGCCGCA 2342
381 LysLeuValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerVal 400
2343 AAGCTGTGCGATTGGGATCAATGCCGTGCCCTACTACCGCGTCTTGAGCTCTCGTC 2402
401 IleProThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThr 420
2403 ATCCCGACCGCGCGATGTTGCTGTGGCAACCGATGCCCTCATGACCGGCTATACC 2462
421 GlyAspPheAspSerValIleAspCysAsnThrCysValThrGlnThrValAspPheSer 440
2463 GCGGACTTCGACTCGGTGATAGCTGCAATACGTGTGTACCCAGACAGTGCATTTAGC 2522
441 LeuAspProThrPheThrIleGluThrIleThrLeuProGlnAspAlaValSerArgThr 460
2523 CTTGACCTACCTTCAACATTGAGACAATCAGCTCCCGCCAGGATGCTGTCTCCGCACT 2582
461 GlnArgArgGlyArgThrGlyArgGlyLysProGlyIleTyrArgPheValAlaProGly 480
2583 CAACGTCCGGGCGAGACTGGCAGGGGAGGAGGACATCTACAGATTGTGGCACCGGGG 2642
481 GluArgProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCys 500
2643 GAGCGCCCTCCGGCATGTTGACTCGCTCTCTGTGAGTGTGTATGACGAGGCTGT 2702
501 AlaTyrTyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThr 520
2703 GCTTGTGATGAGCTCACGCCCGCGAGACTACAGTAGGCTACGAGCGGTATGAAACACC 2762
521 ProGlyLeuProValCysGlnAspHisLeuGluPheTyrGluGlyValPheThrGlyLeu 540
2763 CCGGGGCTTCCGCTGTGTCAGGACCATCTTGAATTTGGGAGGGGCTTTTACAGGCTC 2822

```

```
Qy 541 ThrHisIleAspAlaHisPheLeuSerGlnThrIysGlnSerGlyGluAsnLeuProTyr 560
Db |||||
Db 2823 ACTCATATAGTGGCCACTTCTATCCACACAAGCAGAGTGGGAGAACCTTCCTTAC 2882
Qy 561 LeuValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAsp 580
Db |||||
Db 2883 CTGGTAGCGTACCAAGCCACCGTGTGGCTAGGCTCAAGCCCTCCCCCATCGTGGGAC 2942
Qy 581 GlnMetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeu 600
Db |||||
Db 2943 CAGATGTGGAGTGTGTTGATTCGCTCAAGCCACCTCCATGGGCCAACACCCCTGCTA 3002
Qy 601 TyrArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProValThrLysTyrIle 620
Db |||||
Db 3003 TACAGACTGGGCGGTGTTTCAAGATGAATCACCCTGACGACCCAGTCACCAATACATC 3062
Qy 621 MetThrCysMetSerAlaAspLeuGluValValThrSerThrTrpValLeuValGly 640
Db |||||
Db 3063 ATGACATGCATGTCGGCCGACCTGGAGTCTGTCACGAGCACCTGGGTGCTCGTTGGCGGC 3122
Qy 641 ValLeuAlaAlaLeuAlaAlaTyrCysLeuSerThrGlyCysValValIleValGlyArg 660
Db |||||
Db 3123 GTCCTGGCTGCTTTGGCCGGGTATTGCTGTCAACAGGCTGCGTGTATAGTGGCAGG 3182
Qy 661 IleValLeuSerGlyLysProAlaIleIleProAspArgGluValLeuTyrArgGluPhe 680
Db |||||
Db 3183 GTCGTCTGTTCGGGAAAGCCGGCAATCATCTGACAGGGAAGTCTCTACCGAGAGTTC 3242
Qy 681 AspGluMetGluGluCys 686
Db |||||
Db 3243 GATGAGATGGAAGAGTGC 3260
```

Search completed: September 16, 2004, 00:57:45
Job time : 913 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 15, 2004, 15:46:56 ; Search time 4559 Seconds
(without alignments)
4493.410 Million cell updates/sec

Title: US-09-930-591-2
Perfect score: 3618
Sequence: 1 MAPITAYAQOTRGLGCIIT.....PALIPDREVLRYEPDEMEEC 686

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USFTO_spool/US0930591/runat_15092004.164646.7108/app.query.fasta_1.839
-DB=EST -Qfmt=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US0930591@cgn 1.1 3437 @runat_15092004.164646.7108 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPE=6
-FGAPEXT=7 -YGAPOPE=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:
2: em_esthum:
3: em_estin:
4: em_estm:
5: em_estov:
6: em_estpl:
7: em_estro:
8: em_hc:
9: gb_est1:
10: gb_est2:
11: gb_hc:
12: gb_est3:
13: gb_est4:
14: gb_est5:
15: em_estfun:
16: em_estom:
17: em_gss_hum:
18: em_gss_inv:
19: em_gss_pln:
20: em_gss_vrt:
21: em_gss_fun:
22: em_gss_mam:
23: em_gss_mus:
24: em_gss_pro:
25: em_gss_rod:
26: em_gss_phg:
27: em_gss_vrl:
28: gb_gss1:

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	163	4.5	3291	11	BC046781	BC046781 Mus muscu
2	150.5	4.2	3229	11	AK004665	AK004665 Mus muscu
3	141	3.9	3359	29	AY411078	AY411078 Homo sapi
4	140.5	3.9	822	13	BU054791	BU054791 UT-M-PD0
5	136.5	3.8	3320	11	BC059369	BC059369 Homo sapi
6	133.5	3.7	889	14	CD359697	CD359697 AGENCOURT
7	132.5	3.7	818	13	BU187274	BU187274 AGENCOURT
8	131	3.6	1201	9	AL560974	AL560974 AL560974
9	129	3.6	2972	11	AK031534	AK031534 Mus muscu
10	129	3.6	2984	11	AK031679	AK031679 Mus muscu
11	129	3.6	2986	11	AK028274	AK028274 Mus muscu
12	129	3.6	3956	11	BC043699	BC043699 Mus muscu
13	128.5	3.6	1697	29	AY404177	AY404177 Mus muscu
14	127.5	3.5	691	13	BU054966	BU054966 UT-M-PD0
15	127.5	3.5	738	14	CD240900	CD240900 AGENCOURT
16	127.5	3.5	2976	29	AY400284	AY400284 Homo sapi
17	126	3.5	1283	13	BU0709745	BU0709745 AGENCOURT
18	124.5	3.4	623	14	CF131593	CF131593 UT-HP-FQ0
19	124.5	3.4	2388	29	AY418898	AY418898 Mus muscu
20	124.5	3.4	3633	11	AY383690	AY383690 Rattus no
21	124	3.4	2388	29	AY418896	AY418896 Homo sapi
22	123.5	3.4	1788	11	AY105041	AY105041 Zea mays
23	123	3.4	790	14	CK017540	CK017540 AGENCOURT
24	123	3.4	3461	29	AY398774	AY398774 Mus muscu
25	123	3.4	3802	11	AK004733	AK004733 Mus muscu
26	123	3.4	4327	11	BC058331	BC058331 Mus muscu
27	123	3.4	4640	11	BC062885	BC062885 Mus muscu
28	122	3.4	3879	11	BC028405	BC028405 Homo sapi
29	121.5	3.4	3211	11	BC041392	BC041392 Homo sapi
30	121	3.3	2270	29	AY418897	AY418897 Pan trogl
31	120.5	3.3	478	10	BF725559	BF725559 bx18a02.Y
32	120.5	3.3	694	12	BI088407	BI088407 602852901
33	120.5	3.3	3856	11	AK078552	AK078552 Mus muscu
34	120	3.3	1143	12	BM926541	BM926541 AGENCOURT
35	120	3.3	3494	11	BC013208	BC013208 Homo sapi
36	119	3.3	919	13	EX424517	EX424517 BX424517
37	119	3.3	1339	14	CF753594	CF753594 EST-Conti
38	119	3.3	3679	11	AK084541	AK084541 Mus muscu
39	118.5	3.3	769	9	AU170397	AU170397 AU170397
40	118.5	3.3	1008	12	BQ050646	BQ050646 AGENCOURT
41	117.5	3.2	2031	11	BC038434	BC038434 Homo sapi
42	117.5	3.2	3024	11	AK049321	AK049321 Mus muscu
43	117	3.2	3008	11	AF318356	AF318356 Homo sapi
44	117	3.2	3461	29	AY398772	AY398772 Homo sapi
45	117	3.2	3915	11	BC038417	BC038417 Homo sapi

ALIGNMENTS

RESULT 1
BC046781
LOCUS BC046781
DEFINITION Mus musculus DEAH (Asp-Glu-Ala-His) box polypeptide 16, mRNA (CDNA clone IMAGE:5716182), containing frame-shift errors.
ACCESSION BC046781
VERSION BC046781.1
KEYWORDS HTc.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3291)

Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

TITLE

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA

COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.

FEATURES

source

Location/Qualifiers

1..3359

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

<1..3359

/gene="TIE"

/locus_tag="HCM4104"

ORIGIN

Alignment Scores:

Pred. No.: 0.323 Length: 3359

Matches: 141.00

Percent Similarity: 31.38%

Best Local Similarity: 20.92%

Query Match: 3.90%

DB: 29

US-09-930-591-2 (1-686) x AY411078 (1-3359)

QY 39 ThrAlaAlaGlnThrPheLeuAlaThrCysIleAsnGlyValCysTrpThrValTyrHis 58
 DB 42 ACGAGCCCCAGCGCTCTTCCTGACTTGGGTCTGGGAGGCC----- 86

QY 59 GlyAlaGly-----ThrArg 63
 DB 87 GGGCGGGGAGGGCTCGGACGCTGGGGCCGCCCTGCTGTGGAGAGGACGACCT 146

QY 64 ThrIleAlaSerProLysGlyProValIleGlnMetTyrThrAsnValAspGlnAsp--- 82
 DB 147 ATCGTGCACCCCGCGCGCCACCCCTCGCTCGCGCGCAACGCGTTCGCACACGATC 206

QY 83 ---LeuValGlyTrp-ProAlaPro----- 89
 DB 207 ACGCTTCGGGCTCTCCAGCCCTCGGACCTCGTGGGGCTCTTCCTCGGTGGCGGT 266

QY 90 ---GlnGlyAlaArgSerLeuThrProCysThrCysGlySerAspLeuTyrLe 107
 DB 267 GCTGGGGCGCGCGCGCGCTACCTAGTCACACAGCGCCTGGAGCCACCTGCTT 326

QY 107 uValThrArgHis-----AlaAspValIleProVal----- 117
 DB 327 CGACACAGTCAACACACTGTGAACAAAGGTGACCGCTGTACTTTCTGCACGTGTG 386

QY 118 ---ArgArgGlyAspGlyArgGlySerLeuLeuSerProArgProIle 133
 DB 387 CACAGGAGAGGACAGACAGCGTGTCTGAGAGCAACGGATCTCTTACACCTG 446

QY 133 eSerTyrLeuLysGlySerSerGlyProLeuLeuCysProAlaGlyHisAlaValGle 153
 DB 447 GACTGCATCAAGCCAGGATGGCGGTTC-----TCTGCAGCTCCCAATGTGCAG 500

QY 153 yIlePheArgAlaAlaValCysThrArgGlyValAlaLysAlaValAspPheIleProVa 173
 DB 501 CCACCATCGAGCGGCA-----TCTACAGTGCCT 530

QY 173 lGluSerLeuGluThrThrMetArgSerProValPheSerAspAsn----- 188
 DB 531 TACCTGGAGAGCCAGCCCTCGGACGCGCTCTTTTCGGCTCATCGTGGGGGTGTGGG 590

QY 189 ---SerSerProProAlaValProGlnSerTyrGlnValAlaHisLeu----- 203
 DB 591 GCTGGCGGTGGGGCCAGCTGTACAAAGAGTGCAGGTTGCTCATCATGAGAGTGC 650

QY 204 -----HisAlaProThrGlySer---GlyLysSerThrly 214

DB 651 TGCCACGACCATGACGGGAATGTATGCCCCCTGGCTTCACTGGCACCCGCTGTGAA 710

QY 214 sValProAlaAlaTyrAla-----AlaGlnGlyTyrLy 225
 DB 711 CAGGCTGCAGAGAGGGCGGTTTGGGAGAGCTCCAGGAGCAGTGCAGGAGCATAICA 770

QY 225 sValLeuValLeuAsnProSer-----ValAlaAlaThrMetGlyPheGle 240
 DB 771 GGCTCCCGGGCCTCACCTTCTGCTCCAGACCCCTATGGCTGCTTGTGATCTGCG 830

QY 240 yAlaTyrMetSerLysAlaHisGlyIleAspProAsnIleArgThrGlyValValThrIle 260
 DB 831 TGGAGAGGAGCCAGTGCACAGACCTGTGCCCCCTGTCATTTCGGG----- 879

QY 260 eThrThrGlySerProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCy 280
 DB 879 ----- 879

QY 280 sSerGlyGlyAlaTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSe 300
 DB 880 -----CTGATTGCCAGCTCCAGTCCAGTGTGCAGATGGTGGC 917

QY 300 rIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThrAlaGlyAlaArgLeuThrVa 320
 DB 918 ACTTGTG---ACCGGTTCAAGTGTGTCTGCTGCCCTCTGGGTGCGCATGGAGTGCACGT 974

QY 320 lLeuAla---ThrAlaThrPro----- 326
 DB 975 GAGAGTCAGACCGGATCCCCAGATCTCAACATGGCTCAGAACTGAGATTCACTTA 1034

QY 327 ---ProGlySerValThr-----ValProHisPro----- 335
 DB 1035 GAGACGATGCCCGGATCACTGTGACGTGCAGGGAACCCCTCCCGCTGGGGGACG 1094

QY 336 ---AsnIleGluGluValAlaLeuSerThrThrGlyGluIleProPheTyrGlyLy 353
 DB 1095 ATAGAGCTACCAAGCCAGACGCGACTGTGCTCTGCCACCAAGGCCATTGTGAGCCA 1154

QY 353 sAlaIlePro-----LeuGluAlaIleLysGlyGle 363
 DB 1155 GAGAGACACAGCTGAGTTCGAGGTGCCCGCTTGGTTCTTGGGACAGTGGGTTCGTG 1214

QY 363 yArgHisIleuIlePheCysHisSerLysLysCysAspGluLeuAlaLysLeuVa 383
 DB 1215 CAGTGGCGGTG-----TGTCACATCTGGCGGCCAAGACAGC 1250

QY 383 l-AlaLeu-GlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
 DB 1251 CGGCGCTTCAAGGTCAAT-----GTGAAAGTCCCCCCC 1283

QY 403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThr----- 417
 DB 1284 -----GTGCCCTGCTGCACCTCGGCTCTGACCAAGCAGAGCGCCGACG 1328

QY 418 ---GlyPheThrGlyAsp----- 422
 DB 1329 CTTGTGTCTCCCGCTGGTCTCTCTCTGGGATGGACCCATCTCCACTGTCGCGCTG 1388

QY 423 ---PheAspSerValIleAspCysAsnThrCysValThrGlnThrValAsp 438
 DB 1389 CACTACCGGCCCCAGGACAGTACCATGACCTGGTTCGACCATTTGTG----- 1433

QY 439 PheSerLeuAspProThrPheThrIleGluThrIleThrLeu-----ProGln 454
 DB 1434 -----GTGACCCCACT-----GAGAACTGACGTTAATGAACCTGAGGCCAAG 1478

QY 455 AspAlaValSerArgThrGlnArgGlyArgThrGlyArgGlyLysProGlyIleTyr 474
 DB 1479 ACAGATACAGTGTCTGTGAGCTGAGCCGCGCCAGGGAAGGA----- 1523

QY 475 ArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSer-----SerVa 491
 DB ----- 491


```

QY 279 GlyCysSerGly-----GlyAlaTyrAspIleIleIleCys 290
Db 481 ATGCTGGTCAGGAATGATGGTTGATCCGTTGTTAAACAAATATAGTGTATCATGCTG 540
QY 291 AspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGln 310
Db 541 GATGAGCCACGAGGAGACCTTGATACACTGCATTCGCAATGGCTTGCTATAAAGATT 600
QY 311 AlaGluThrAlaGlyAlaArgLeuThrValLeuAlaThrAlaThrProGlySerVal 330
Db 601 CAGAAAAGCGAGGGATCTTCGATTGATGTAGCTTCAGCCACTCTGGATGCAGACAAA 660
QY 331 ThrValProHisProAsnIleGluGlu-----ValAla 341
Db 661 TTCGGGATTTCTTTAATCAAAATGAAACCAAGTATCCAGCAAGGATACATGTGTGATC 720
QY 342 LeuSerThrThrGlyGlu-----IlePro 349
Db 721 CTTACAGTGAAGGAGAAACATTTCGCTGGATATCTTTATATACAAAGTCTGTTCCTCA 780
QY 350 PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyGly 363
Db 781 GATTATATCAATCAACTGTCGAACTGTGGTGTGAATTCACACAGAGGAGACGGA 840
QY 364 ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaIleVal 383
Db 841 GAGCTTTTAGCATTTCTTACTGCGCAGGAAGGTAGAAACTGTGTGTCGATGCTCATC 900
QY 384 -----AlaLeuGly-----ValAsnAlaValAlaIleTyr 392
Db 901 GAGCAGCTCGAGCACTAGCTCGCACTGGATGAAGACACCTCCGAGTTCCTCCCATG 960
QY 393 TyrArgGlyLeuAsp-----ValSerValIleProThrSerGly 405
Db 961 TATCGAGGACTGCTCTCTTTGAGCAATGAAAGTGTGTGAAAGGGTGTCAAGCATGTC 1020
QY 406 AspValValValAlaThrAspAlaLeuMetThrGlyPheThr---GlyAspPheAsp 424
Db 1021 AGAAGGTGATAGTGGCCACCAATGTGGCAGAAACCTCTATCAATCATGACGGCATGTG 1080
QY 425 SerValIleAspCysAsnThrCysValThrGlnThrValAspPheSerLeuAspProThr 444
Db 1081 TATGTGATCGACTGTGCTTTGTGAAACTCCGA-----GCTTACAAATCCAGG 1128
QY 445 PheThrIleGlu---ThrIleThrLeuProGlnAspAlaValSerArgThrGlnArg 463
Db 1129 ACAGCTATTGAATGCTTTGGTGGTGCCAGTCTCCCAAGCATCACTAATCAGCGACGA 1188
QY 464 GlyArgThrGlyArgGlyLysProGly---IleTyrArgPheValAlaProGlyGluArg 482
Db 1189 GGAGTGGTGTCTGATGCTCGCGGAAATGTTATCGCTTTATACA-----GAGGAA 1242
QY 483 ProSerGlyMetPheAspSerSerValLeuCysGluCysTyrAspAlaGlyCysAlaTrp 502
Db 1243 GCCTTTGACAAGTTGCCCTCATGCTACGGTTCCTGAGATGCGCGGTAGTAATTG----- 1296
QY 503 TyrGluLeuThrProAlaGluThrThrValArgLeuArgAlaTyrMetAsnThrProGly 522
Db 1297 -----GCACTGTCTATCTCGACGCTGAAAGCA----- 1323
QY 523 LeuProValCysGlnAspHisLeuGluPheTrpGluGlyValPheThrGlyLeuThrHis 542
Db 1324 -----CTAGGAATTGCAAT 1338
QY 543 Ile---AspAlaHisPheLeuSerGlnThrLysGlnSerGlyGluAsnLeuProTyrLeu 561
Db 1339 GTCTCAGGTTCCACTTCATGTC----- 1362
QY 562 ValAlaTyrGlnAlaThrValCysAlaArgAlaGlnAlaProProSerTrpAspGln 581
Db 1363 -----CCCCCTCCAGCA---CAGTCG 1380

```

```

QY 582 MetTrpLysCysLeuIleArgLeuLysProThrLeuHisGlyProThrProLeuLeuTyr 601
Db 1381 ATGGTTCAAGCCTTG-----GAGTTACTGTAT 1407
QY 602 ArgLeuGlyAlaValGlnAsnGluValThrLeuThrHisProVal 616
Db 1408 GCTCTGGAGGCTCTGGCAAGAACTGTGCGCTAACTGAACCGCTT 1452

```

RESULT 6
CD359697
LOCUS
DEFINITION
IMAGE:30388971 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD359697 889 bp mRNA linear EST 29-MAY-2003
AGENCOURT 14284682 NIH_MGC_180 Homo sapiens cDNA clone
IMAGE:30388971 5', mRNA sequence.
CD359697
EST
CD359697.1 GI:31131108
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM461 row: 1 column: 04
High quality sequence start: 22
High quality sequence stop: 653.
Location/Qualifiers
1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30388971"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH_MGC_180"
/notes="Organ: Testis; Vector: pCMV-SPORT6.1; Site: 1. Not I;
Site_2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

FEATURES
source
1..889
Alignment Scores:
Pred. No.: 0.157 Length: 889
Score: 133.50 Matches: 61
Percent Similarity: 40.24% Conservative: 40
Best Local Similarity: 24.30% Mismatches: 95
Query Match: 3.69% Indels: 55
DB: 14 Gaps: 11
US-09-930-591-2 (1-686) x CD359697 (1-889)

QY 183 ProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrGlnValAlaHis 202
Db 44 CCGGTGTTCCCATTCGAGAGAGAGCTCTGCTGCTCTATTGCAATCACCAGTCTCTCATC 103
QY 203 LeuHisAlaProThrGlySerGlyLysSerThrLysValPro-----Ala 217
Db 104 ATTGAAGCGGAGAGAGGCTCAGGGAAGACACCAGATCCCGCAGTATCTCTTTGAGGAG 163

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sui, N., Ishii, Y., Nakamura, S., Hazama, M., Nishise, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
PUBMED
11076861

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature.409, 685-690 (2001)

JOURNAL

REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

JOURNAL

REFERENCE

6 (bases 1 to 2972)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

TITLE

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

FEATURES

source

1. .2972
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:6030448F20"
/db_xref="MGI:2394772"
/db_xref="taxon:10090"
/clone="6030448F20"
/sex="male"
/issue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="13 days embryo"
17. .2641
/notes="unnamed protein product; hypothetical p-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOP|52540, evidence: SCOP) putative"
/codon_start=1
/protein_id="BAC27438.1"
/db_xref="GI:26327389"

CDS

```

/translation="MAAGRRVGPSPRPRTMAPWKRLKRLKRRKTCASQGRSDSDGGE
FEIQAEDDARARKLGPRLSPPTSCVSDVEPTREMVRAQNKKKKSGQFQSMGL
SYVPFKGIMKGYKVPPIQKRTIPVLDGKVAVAMATGSKTACTFLPWFELKAR
SACTGARALILSPRELALQMTKTELKFTGLKTAJLIGDKWEDQFAALHFNKPD
IATPGRLVHVAVEMNLKQSVYVDFEADRLFEMGFACQLOELIIGLPGHQTULF
SATIPKLVEPAQAGLTPVLRLDVSQKNEQLKTFLLVREDTKAVALLYLQNVV
RPQOQTVVATKHAHYLTLLMGQVSCAHYISALDQTKARKINLAKFTNKKCSSTLI
VTDLAAGLDIPLDNLNINVSFPKGLFLHRGVRVARAGSGTAYSLVAFVPEYLL
DLHLFLGRSVTLARCEEPSVADAGVLDGVRVQSVDDSDSLQTMASGLDLQGL
LHRVANNAQQOYVRSRPAPESEIKRAKELDLAELGLHPLFSSCFEGEQLRLVDS
IKNVRTTTFEINASGKDPSSOMRAKQRDRKAVASFOORROEROGEPADPAQRE
LPOREEREMVETVGVTEVVQKRPREGPSQGAKRREMETRQDQREYVYVRPKDED
SERGLSVSGACGAEQOVAGAVLDIMGEAQNMRSGQQLKWDKRRKRFVQCSQDEK
KKIKTSERFISSYKLDYQKWKQKIDDRSEEGPSNQRGPRRGGRRSQSG
TSQPRASSVPAGRMRSSELTKEQLKQRRQAKQKORFLQRRGLKQLSARNRRRAQELRQ
GAFGRGAPSRKGRKRM"
2954. .2959
/note="putative"
2972
/note="putative"
polya_signal
polya_site
ORIGIN
Alignment Scores:
Pred. No.:      3.47      Length:      2972
Score:          129.00    Matches:     134
Percent Similarity: 31.24% Conservative: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match:      3.57% Indels:      201
DB:               11      Gaps:       29
US-09-930-591-2 (1-686) x AK031534 (1-2972)
QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114
Db 90 TCGGAAACGCCGAACTGGGCGCTTCCCAAGCCGCGACAGCAGCTGGATGACGCGAGT 149
QY 115 IleProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
Db 150 TCGAGATCCAGCGGAGGATGACGCCGCGGCGAGGAAGCTGGCGCCTGGCAGACCTTGC 209
QY 132 Pro-IleSerTyrLeuIys-----GlySerSerGlyGlyProLeuLe 145
Db 210 CCTCATTTTCCATCTCAGCTGCGTATCAGATGTGGAGCCGACACATCGGAGAGTGTGTC 269
QY 145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
Db 270 GAGCCAGAACACAAAGAAAGAGAGTCTGGAGGCTTCCAGTCCATGGGCTGAGTTACC 329
QY 160 -CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThr 179
Db 330 CTGTGTTTCAGGGGATCATGAAAAAAGGGCTACAAAGTGGCGGACGCCCATCCAGAGGAAGA 389
QY 179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG 199
Db 390 CCATC-----CCCGTGATCTTGGATGGCAAGGATGTGGTGGCCATG----- 430
QY 199 lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
Db 431 -----GCCCGACAGCGCAGTGGCAAGAGCGGCGCTGCTTCTCTCTCCCGCA 473
QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn 231
Db 474 TGTTTGAGCGGCTGAAGCGCAGCGAGTGCAGACGCGGGGCTCGAGCGCTCATCTCTCTCAC 533
QY 231 rSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248
Db 534 CCACCCGGAGCTGGCCCTCGACACCATGAAGTTCCTAAAGAGCTAGGCAAGTT-CACC 592
QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlyS 264
Db 593 GGCCTCAAGACTGCGCTTGATCTGCTGGTGGAGACAAAATGGAAGACCACTTTCAGCCCTG 652
QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGlyA 284

```

```

Db      653  CACGAGAACCTGACATAATCAATTGCCACCCCTGGGCGTCTGGTGCATGT-----GGCT 706
Qy      284  laTyraSpIleIleLeuCyAspGluCyHisSerThrAspAlaThrSerIleLeuGlyI 304
Db      707  GTGGAGATGAACCTTGAAGCTGCAGAGT-----GGAGTATGTGGTG 748
Qy      304  le----- 304
Db      749  TTCGATGACGACACAGGCTCTTTGAATGGGCTTTGCTGAGCAGCTACAGGAGATCATA 808
Qy      305  -----GlyThrValLeuAsp---GlnAlaGluThrAlaG 315
Db      809  GGCGCGCTTCTCGGGGGCCACGACGGTCTGCTCTCAGCTACACTGCCCAAGTGGCTG 868
Qy      315  lYAlaArgLeuThr-----ValLeuAlaThrAlaThrProGlySerValThrV 332
Db      869  GTGGAATTCGACAGGAGGCTCAGAGCCGCTGCTCATCGCTGGAGTGGAGTACTCC 928
Qy      332  alProHisProAsnIleGluLeuValAlaLeuSerThrThrGlyGluLeuPro----- 349
Db      929  AAGCTCAATCAGCAGCTCAAGACCTC-----CTTCCTCTGTG 967
Qy      350  -----PheTyrglyLeuAlaIleProLeuAlaIleLys-----GlyG 363
Db      968  CGGAGACACCAAGGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1027
Qy      363  ly-ArgHisLeuIlePheCyHisSerLysLysLysCysAspGluLeuAlaLysLeu 382
Db      1028  GACCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1087
Qy      383  ValAlaLeuGlyValAsnAlaValAlaTyrrArgGlyLeuAspValSerValIlePro 402
Db      1088  ATGGCCAGGCTGTGATGCTGGCCCATCTATAGTGCCTGTGACACGACG----- 1138
Qy      403  ThrSerGlyAspValValValAlaValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
Db      1139  -----GCCGCGCAAGTCAACTGTGCGCAAGTTCACACACA 1174
Qy      423  PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
Db      1175  -----AAATGTTCCACCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1219
Qy      442  AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db      1220  GACATCCACTGTGGACAACGTCATCACTACAGCTCCCTCCCAAGGCAAGCTCTTC 1279
Qy      457  -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIle 473
Db      1280  CTGCACCGAGTGGCGCGTGTGGCCGAGCAGGCGGAAGTGGCACAGCC----- 1327
Qy      474  TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerSerValLeuCy 493
Db      1328  TATTCITTTGGTGGCCCGACAGAGTCCCTTACC----- 1361
Qy      493  sGluCysTyraSpAlaGlyCysAlaTptTyrgluLeuThrProAlaGluThrThrValAr 513
Db      1362  ----TGCTTGACCTACACCTGTTCTCGGCGGCTGTGTACCCCTGGCGCTCTTGAG 1417
Qy      513  g-----LeuArgAlaTyrrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db      1418  GAGCTTTCAGTGGCAGATGCGGTGGCAGGG-----ACGAGGTG 1456
Qy      532  eTrpGluGly----- 535
Db      1457  CTGGTGGTGGTGGCCCGCAGGTGTAGTGTGATGATGAGGACGACGACGCTGCGAGTGGCCATG 1516
Qy      536  -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db      1517  GGGGCATCCCTGGATCTTCAGGGCTGCACC----- 1547
Qy      551  rLysGlnSerGlyGluAsnLeuProTyrrLeuValAlaTyrrGlnAlaThrValCysAlaAr 571

```

```

Db      1548  -----GGTGGCCCAACAACGCTCAGCAGCAGTATGTGGCTCA 1585
Qy      571  gAlaGlnAlaPro-----ProSerTrpAsp---GlnMetTr 583
Db      1586  CGGCAGCGCCCTCGCTGCTGAGTCCATCAGAGAGCCCAAGGAGCTGGACCTGGCAGAGCTG 1645
Qy      583  pLysCys 585
Db      1646  GGCTTGC 1652

RESULT 10
LOCUS   AK031679
DEFINITION Mus musculus 13 days embryo male testis cDNA, RIKEN full-length
           enriched library, clone:6030486C03 product:hypothetical P-loop
           containing nucleotide triphosphate hydrolases structure containing
           protein, full insert sequence.
ACCESSION AK031679
VERSION   AK031679.1 GI:26327530
KEYWORDS  HTC; CAP trapper.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
           Eukaryota; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
  1 Carninci,P. and Hayashizaki,Y.
    High-efficiency full-length cDNA cloning
    Meth. Enzymol. 303, 19-44 (1999)
  2 99279253
    10349636
  3 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
    Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
    Normalization and subtraction of cap-trapper-selected cDNAs to
    prepare full-length cDNA libraries for rapid discovery of new genes
    Genome Res. 10 (10), 1617-1630 (2000)
  4 20499374
    11042159
  5 Shibata,K., Itoh,M., Aizawa,K., Nagasaka,S., Sasaki,N., Carninci,P.,
    Konno,H., Akiyama,J., Nishi,K., Kitaura,T., Tashiro,H., Itoh,M.,
    Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishize,T., Harada,A.,
    Yamamoto,K., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
    Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Matsuoka,M.,
    Yoshida,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
    Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
    RIKEN integrated sequence analysis (RISA) system--384-format
    sequencing pipeline with 384 multicapillary sequencer
    Genome Res. 10 (11), 1757-1771 (2000)
  6 20530913
    11076861
  7 The RIKEN Genome Exploration Research Group Phase II Team and the
    FANTOM Consortium.
    Functional annotation of a full-length mouse cDNA collection
    Nature 409, 685-690 (2001)
  8 The FANTOM Consortium and the RIKEN Genome Exploration Research
    Group Phase I & II Team.
    Analysis of the mouse transcriptome based on functional annotation
    of 60,770 full-length cDNAs
    Nature 420, 563-573 (2002)
  9 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
    Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
    Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
    Hori,F., Imocani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
    Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
    Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
    Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
    Okazaki,Y., Saito,R., Saichoh,H., Sakai,C., Sakai,K., Sakazume,N.,
    Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
    Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,

```

Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gs.c.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1 . 2984
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:6030486C03"
/db_xref="MGI:2334945"
/db_xref="taxon:10090"
/clones="6030486C03"
/sex="male"
/tissue_type="testis"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stages="13 days embryo"
25 . 2754
/note="unnamed protein product; hypothetical P-loop containing nucleotide triphosphate hydrolases structure containing protein (SCOPE|52540, evidence: SCOP) putative"
CDSD
/codon_start=1
/protein_id="BAC27509.1"
/db_xref="GI:26327531"
/translat="MAAGRVGGPPSPRPTMAPKKKRLKRRTGASQGRSDSDGGE FEIQAEDDARAKLPGORALPSPTSECVDSVPDTREMVRAQNKKKKSGGFQSMGL SYFGIMKKGVYVPIQRKTIPTVLGDKVVMARTGSQGTACFLLPMEERLAR SAQTGARALLSPTRELALOTMAKTKELGKFTGLKTALLIGDKMEDQPAALHENDFI IATPGRLVHAVEMVLKLSQSVYVDFEADRLFEMGPAPQLQEIIGRUPGGHTVILF SATPLKLIVFPARAGLTETPVLLIDVDKNEQLKTSFLVRVDTRKAVALLYLLQNV RPQDQVFVTADIBAHYLTLEMLQGQVSCAHIYSALDOTARKINIAKTFTHNKCSTLI VTDLAQGLDIPLDVINFSYFAKGKLFHLRVVARAGRSCTAYSLVAPDSLQGL DLHLFLGRSYTLARPCEPESVADAVRGDVLGRVQSVDDESSITGMAGSLDQGL LHRVANAOQQVYRSPAPSPESIIRAKELDIAEGLHPLFSSCFEGELQRLRVDS IKNYRRTTIFETNASGDSPSSOMRAKORDRKVASFOQRQEQQEQGPAPAPORE LPQEEEMEVTVEGVTEVTVGCKRPGPSQCAKRMETRDQDEFVTPVPKPQDF SERGLSVSGAGGAPEQAVAGAJDLMDGNQSKMRMTKWDRKKKRFVGSGSQED KTKITSGRISFYKRDLYKQKQKIDRRSEEGPSNQRPGRGKRGKRSQG TSQPSPASSPMRSELKTKQIILKRRQGSAAALPAARGPEAFSTOPTSPSGAAP GRWPDCSLQEGDEEKVESQTPQWGLVAPGDVDRVHVLLDPSTGAAL"

polyA_signal
polyA_site
ORIGIN

Alignment Scores:	
Pred. No.:	3 5 Length: 2984
Score:	129, 00 Matches: 134
Percent Similarity:	31.24% Conservatise: 55
Best Local Similarity:	22.15% Mismatch: 216
Query Match:	3.57% Indels: 201
DB:	11 Gaps: 29

US-09-930-591-2 (1-686) x AK031679 (1-2984)

QY 100 CysglySerAspLeuTyIeu-----ValThrArgHisAlaaspval 114
|||||:.....:||||
|||

polyA_signal putative"
2960..2965 /note="putative"
polyA_site 2986
/note="putative"
ORIGIN

Alignment Scores:

Pred. No.: 3.5 Length: 2986
Score: 129.00 Matches: 134
Percent Similarity: 31.24% Conservativity: 55
Best Local Similarity: 22.15% Mismatches: 216
Query Match: 3.57% Indels: 201
DB: 11 Gaps: 29

US-09-930-591-2 (1-686) x AK028274 (1-2986)

```

QY 100 CysGlySerSerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114
DB 94 TCCGAAACGCCGAACCTGGGCTTCCCAAGCCGCGACAGCGACTCGGATGACGGCGAGT 153
QY 115 IleProValArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
DB 154 TCGAGATCCAGCGGAGGATGACGCCCGGCGAGGAGCTGGGCCCTGGCAGAGCCTTGC 213
QY 132 Pro-IleSerTyrLeuLys-----GlySerSerGlyGlyProLeuLe 145
DB 214 CCTATTCTTACCTCAGAGTGCGTATPCAGATGTGGAGCCCGACACTCTCGGAGATGGTGC 273
QY 145 uCysProAlaGlyHisAlaValGlyIlePheArgAlaAlaVal----- 159
DB 274 GAGCCAGACACAGAAAGAGAGTCTGGAGCTTCCAGTCCATGGGCGCTGAGTTACC 333
QY 160 -CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGlySerLeuGluThrt 179
DB 334 CTGTGTTCAAGGGGATCATGAAAGGGCTACAAAGGTGCGCGCCGCTCCACAGGAGGA 393
QY 179 hrMetArgSerProValPheSerAspAsnSerSerProProAlaValProGlnSerTyrG 199
DB 394 CCATC-----CCCGTATCTTGATGGCAAGGATGTGGGCCATG----- 434
QY 199 lnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
DB 435 -----GCCCGACAGGACAGTGTGCAAGAGCGGCTGCTTCTCTCCCGA 477
QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsnP 231
DB 478 TGTGTAGCGGCTGAAGGCGACGAGTGCACAGACGGGGCTCGAGCCCTCATCTCTCAC 537
QY 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisG 248
DB 538 CCACCCGGAGCTGGCCCTGCACACCATGAGTTCAATAAGGCTAGGCALAGTT-CACC 596
QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrThrGlys 264
DB 597 GGCTCAAGACTGCTCTGTATCTCTGGTGGGACAAAAATGGAACACAGTTCGAGCCCTG 656
QY 264 exProlIleThrTyrSerThrTyrGlyPheLeuAlaAspGlyGlyCysSerGlyGlyA 284
DB 657 CAGAGAACCTGCACATAATCATGTCACCCCTGGCGCTCTGTGTCATGT-----GGCT 710
QY 284 laTyrAspIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304
DB 711 GTGAGATGAACCTGAAGCTGCAGAGTGT-----GGAGTATGTGGTG 752
QY 304 le----- 304
DB 753 TTCGATGAACGACAGAGGCTCTTTGAAATGGGCTTGTCTGACGAGCTACAGGAGATCATA 812
QY 305 -----GlyThrValLeuAsp-----GlnAlaGluThrAlaG 315
DB 813 GGCGGCTTCTGGGGCCACAGAGCGTGTCTGTCTCTAGCTACACTGCCGCCAGCTGCTG 872

```

```

QY 315 lyAlaArgLeuThr-----ValLeuAlaThrAlaThrProProGlySerValThrV 332
DB 873 GTGGAATTTGCACGGCGAGCCTCACAGAGCCCGTCTCATCCGCTGACGTAGACTCC 932
QY 332 alProHisProAsnIleGluValAlaLeuSerThrThrGlyGluIlePro----- 349
DB 933 AAGCTCAATGACGAGCTCAAGACCTC-----CTTCTCTCTGTG 971
QY 350 -----PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyG 363
DB 972 CGGAAGACACCAAGGCTGCTGTCTCTCTACCTGCTGAGATGTCTGCGGCCAG 1031
QY 363 ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaAlaLysLeu 382
DB 1032 GACCAGCTGTGCTGTTTCTAGCCACAAAGACCATGCGAGTACTCTACAGTTGCTG 1091
QY 383 ValAlaLeuGlyValAsnAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
DB 1092 ATGGGCCAGGGTGTGAGTTGCGCCCATCTATAGTCTTGGACACGAG----- 1142
QY 403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
DB 1143 -----GCCGCAAGATCAACTTGGCCCAAGTTACACACAAAC 1178
QY 423 PheAspSerValIleAspCysAsnThr---CysValThrGlnThrValAspPheSerLeu 441
DB 1179 -----AAATGTCCACCTCATCTGCTGACTGACCTGCGCCCGCGGCGCTG 1223
QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
DB 1224 GATATCCCACTGCTCGCAACAGCTCATCACTACAGCTTCCCTGCGCAAGGCAAGCTTTC 1283
QY 457 -----ValSerArgThrGlnArgArgGlyArgThrGlyArgGlyLysProGlyIle 473
DB 1284 CTGCACCGAGTGGGCGGTGTGCGCGAGCAGCGCAAGTGGCAGACGC----- 1331
QY 474 TyrArgPheValAlaProGlyGlyGlu-ArgProSerGlyMetPheAspSerSerValLeuCy 493
DB 1332 TATCTTTGTGGTGGCCCGCAGAGGAGTCCCTTACC----- 1365
QY 493 sGluCysTyrAspAlaGlyCysAlaTyrTyrGluLeuThrProAlaGluThrValThr 513
DB 1366 -----TGCTTGACCTACACCTGTCTGGCGCGCTGTGACCCCTGGCGCTCTGTGAG 1421
QY 513 g-----LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
DB 1422 GAGCCTTCAGTGGCAGATGCGGTTGGCAGGG-----ACGAGGTG 1460
QY 532 sTyrGluGly----- 535
DB 1461 CTGGTCCGCTGCCCGAGGTGTAGTGATGATGAGGACAGCGCTGCAGACTGCCCATG 1520
QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
DB 1521 GGGGCATCCCTGGATCTTTCAGGGCTGCACC----- 1551
QY 551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
DB 1552 -----GGGTGGCCCAACAGCTCAGCAGCAGTGTGACCTGGACGAGCTG 1589
QY 571 gAlaGlnAlaPro-----ProProSerTyrAsp---GlnMetTr 583
DB 1590 CGGCCAGCGCCCTCGCTGAGTCCATCAAGAGAGCCAGGAGCTGGACCTGGCAGAGCTG 1649
QY 583 rLysCys 585
DB 1650 GGCTTG 1656

```

RESULT 12

BC043699

LOCUS

DEFINITION

BC043699 3956 bp mRNA linear HTC 13-JAN-2003
Mus musculus, Similar to RIKEN cDNA 2410015A15 gene, clone
IMAGE:5344158, mRNA.

ACCESSION BC043699
 VERSION BC043699.1 GI:27696772
 KEYWORDS HTC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3956)
 Strausberg, R.
 Direct Submission
 Submitted (10-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 info@bcgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Leticia Heiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 86 Row: 1 Column: 19
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis
 This clone has the following problem: frame shifted.
 Location/Qualifiers
 1..3956
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVS/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5344158"
 /tissue_type="mammary tumor, C3(1)-Tag model. Infiltrating ductal carcinoma, 5 month old virgin mouse."
 /clone_lib="NCI CGAP Mamms"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN
 Alignment Scores:
 Pred. No.: 5,74 Length: 3956
 Score: 129.00 Matches: 134
 Percent Similarity: 31.24% Conservative: 55
 Best Local Similarity: 22.15% Mismatches: 216
 Query Match: 3.57% Indels: 201
 DB: 11 Gaps: 29

US-09-930-591-2 (1-686) x BC043699 (1-3956)

QY 100 CysGlySerAspLeuTyrLeu-----ValThrArgHisAlaAspVal 114
 DB 80 TCGGAAGACCGCGAGTATCCAGCGCGCGAGAGTGGCGGATGACGGCGAT 139
 QY 115 IleProValArgArgArg-----GlyAspGlyArgGlySerLeuLeuSerProArg 131
 DB 140 TCGAGATCCAGCGCGAGGATGACGCCCGCGGAGAGTGGCGGCTGGCAGAGCCTTC 199
 QY 132 Pro-IleSerTyrLeuLeu-----GlySerSerGlyGlyProLeuLe 145
 DB 1155 -----AAATGTTCCACCTCATCGTACTGACTGAGCTGGCGCGCGGCGCTG 1209

Db 200 CCTCATTTCTACCTCAGAGTGGCGTATCAGATGTGGAGCCCGACACTCGGAGATGGTGC 259
 QY 145 uCysProAlaGlyHisAlaValAlaGlyIlePheArgAlaAlaVal----- 159
 DB 260 GAGCCAGACCAAGAAAAGAAAGTCTGGAGCTTCCAGTCCATGGGCTGAGTTACC 319
 QY 160 CysThr-ArgGlyValAlaLysAlaValAspPheIleProValGluSerLeuGluThrT 179
 DB 320 CTGTGTTCAAGGGGATCATGAAAGAGGCTACAAGTGGCGAGCCCATCCAGAGAGA 379
 QY 179 hrMetArgSerProValPheSerAspAsnSerSerProAlaValProGlnSerTyrG 199
 DB 380 CCATC-----CCCGTATCTTGGATGGCAAGGATGGTGGCCATG----- 420
 QY 199 InValAlaHisLeuHisAlaProThrGlySerGlyLysSerThr----- 213
 DB 421 -----GCCCGGACAGCGAGTGGCAAGCGGCGCTGCTTCTCTCCCGA 463
 QY 214 -----LysValProAlaAlaTyrAlaAlaGlnGlyTyrLysValLeuValLeuAsn 231
 DB 464 TGTGAGCGGCTGAGGCGAGTGCACAGACGGGGCTCGAGCCCTCATCTCTCAC 523
 QY 231 roSer-----ValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis 248
 DB 524 CACCCGGGAGCTGGCCCTGCAGACCATGAAGTTCACTAAGAGCTAGGCAAGTT-CACC 582
 QY 248 ly-----IleAspProAsnIleArgThrGlyValArgThrIleThrGlyS 264
 DB 583 GGCCTCAAGACTGCTTGTCTCTGGTGGAGACAAAATGAAGACAGTTCGAGCCCTG 642
 QY 264 erProIleThrTyrSerThrTyrGlyLysPheLeuAlaAspGlyGlyCysSerGlyGly 284
 DB 643 CAGGAGAACCTGACATAATCATTCGCCACCCCTGGCGCTGGTGCATGT-----GGCT 696
 QY 284 laTyrAspIleIleIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyI 304
 DB 697 GTGGAGATGAACCTGAAGTGCAGAGTGT-----GGAGTATGGGTG 738
 QY 304 le----- 304
 DB 739 TTGATGAGCAGACAGCTCTTTGAAATGGGCTTGTGAGCAGCTACAGAGATCAT 798
 QY 305 -----GlyThrValLeuAsp-----GlnAlaGluThrAlaG 315
 DB 799 GGCGCGCTTCTGGGGGCCACAGACGGTGTCTTCTCAGCTACACTGCCAAGTGTCTG 858
 QY 315 lyAlaArgLeuThr-----ValLeuAlaThrAlaThrProGlySerValThrV 332
 DB 859 GTGGAATTTGACGGGCGAGCCCTCAGAGACCGGTGTCTATCGCGCTGAGTACTTC 918
 QY 332 alProHisProAsnIleGluGluAlaLeuSerThrThrGlyGluIlePro----- 349
 DB 919 AAGTCAATGAGCGCTCAAGACCTC-----CTTCTCTCTTGTG 957
 QY 350 -----PheTyrGlyLysAlaIleProLeuGluAlaIleLys-----GlyG 363
 DB 958 CGCAAGACACCAAGGCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1017
 QY 363 ly-ArgHisLeuIlePheCysHisSerLysLysCysAspGluLeuAlaLysLeu 382
 DB 1018 GACCAGACTGTGTGTTGTAGCCACAAAGCACCATGGCGGTACTCTCAGAGTTGCTG 1077
 QY 383 ValAlaLeuGlyValAlaAlaValAlaTyrTyrArgGlyLeuAspValSerValIlePro 402
 DB 1078 ATGGCCAGAGGTGTGAGTTGGCGCCCATCTATAGTGCCTTGGACCAAGTTCACACAC 1128
 QY 403 ThrSerGlyAspValValValAlaThrAspAlaLeuMetThrGlyPheThrGlyAsp 422
 DB 1129 -----GCCCGCAAGATCACTTGGCCCAAGTTCACACACAC 1164
 QY 423 PheAspSerValIleAspCysAsnThr-----CysValThrGlnThrValAspPheSerLeu 441
 DB 1165 -----AAATGTTCCACCTCATCGTACTGACTGAGCTGGCGCGCGGCGCTG 1209

```

QY 442 AspProThrPheThrIleGluThrIleThrLeuProGlnAspAla----- 456
Db 1210 GACATCCCACTGTGGCAACAGTCATCAACTACAGCTTCCTGCCAAGGCCAAGCTCTTC 1269
QY 457 -----ValSerArgThrGlnArgGlyValArgThrGlyArgGlyLysProGlyIle 473
Db 1270 CTCACCGAGTGGCGTGTGGCCCGCAGCGGCCGGAAGTGGCACACC----- 1317
QY 474 TyrArgPheValAlaProGlyGlu-ArgProSerGlyMetPheAspSerValLeuCy 493
Db 1318 TATTCTTGTGGCCCGCAGCAGAGTCCCTACC----- 1351
QY 493 sGluCysTyrAspAlaGlyCysAlaTrpTyrGluLeuThrProAlaGluThrThValAr 513
Db 1352 ----TGCTTGACCTACACTGTTCCTGGCCCGCTCTGTACCCCTGGCCCGCTCTGTGTAG 1407
QY 513 g----LeuArgAlaTyrMetAsnThrProGlyLeuProValCysGlnAspHisLeuGluPh 532
Db 1408 GAGCCTTCAGTGGCAGATCGGTGGCAGGG-----ACGGAGTG 1446
QY 532 eTrpGluGly----- 535
Db 1447 CTGGGTGGGTGCCCGCAGAGTGATGTGATGATGAGGACAGCAGCCTGCAGACTGCCATG 1506
QY 536 -----ValPheThrGlyLeuThrHisIleAspAlaHisPheLeuSerGlnTh 551
Db 1507 GGGGATCCCTGATCTTACGGCCCTGCACC----- 1537
QY 551 rLysGlnSerGlyGluAsnLeuProTyrLeuValAlaTyrGlnAlaThrValCysAlaAr 571
Db 1538 -----CGGTGGCCACACAGCCTCAGCAGCAGTATGTGGCTCA 1575
QY 571 galaginAlaPro-----ProProSerTrpAsp---GlnMetTr 583
Db 1576 CGCCAGCGCCCTCGCTAGTCCATCAGAGAGCCAGAGCTGACCTGGCAGAGCTG 1635
QY 583 physCys 585
Db 1636 GGCTTGC 1642

RESULT 13
AY404177
LOCUS
DEFINITION
1697 bp DNA linear GSS 15-DEC-2003
Mus musculus DDX34 gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
AY404177
VERSION
GI:39760154
KEYWORDS
GSS.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1697)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED
14671302
REFERENCE
2 (bases 1 to 1697)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
JOURNAL
AUTHORS
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
COMMENT
Location/Qualifiers
FEATURES

```

```

source
1..1697
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1697
/gene="DDX34"
/locus_tag="HCM1803"

gene

ORIGIN

Alignment Scores:
Pred. No.: 1.44 Length: 1697
Score: 128.50 Matches: 134
Percent Similarity: 35.56% Conservative: 68
Best Local Similarity: 23.59% Mismatches: 221
Query Match: 3.55% Indels: 148
DB: 29 Gaps: 30

US-09-930-591-2 (1-686) x AY404177 (1-1697)
QY 52 ValCysTrpThrValTyrHisGlyValAlaGlyThrArgThrIleAlaSerProLysGlyPro 71
Db 116 GTCTCTTGGAAAGATGATTCTTCCGCG---ATGAGGATTATATCC-----GTC 160
QY 72 ValIleGln-----MetTyrThrAsnValAspGlnAspLeuVal 84
Db 161 GTGGTTCAGAGGAGTGTCCAGAAAGTTCTGGGCTTCTTTGAACGCTGCAGAGATTCCAGC 220
QY 85 GlyTrpProAlaProGlnGlyValaArgSerLeuThrProCysThrCysGlySerSerAsp 104
Db 221 ACCTCAAGACCTCCACGAAGAAGAGAGGACCCCTGGCATGCCCAACATGGCATCGCTG 280
QY 105 Leu-----TyrLeu---ValThrArgHisAlaAspValIleProValArgArg 120
Db 281 CTTTAGCTGACCTACCTCTCACTTACGACCCACGCT----- 316
QY 121 GlyAspGlyArgGlySerLeuLeuSerProArgProIleSerTyrLeuLysGlySerSer 140
Db 317 ---ACCGCATCAACCTCTCCATTCTCAGCCAGACACATCGGGGTCGTATGGGCCCGGCA 373
QY 141 GlyGly-----ProLeuLeuCysProAlaGlyHisAlaValGlyIlePheArgAlaAla 158
Db 374 GGGGGTCCCCCGCGAGAGAGTTCGAGTCC---GCCGTGCATTGCTTCACTACCTAG 430
QY 159 ValCysThrArgGlyValAlaLysAlaValAspPheIleProValGluSerLeu---Glu 177
Db 431 ACTTCCAGCAGA-----AGCAGGCTTTTGGAGGCTCGCCAAACTGCAGCGTGAGCAG 484
QY 178 Thr-ThrMetArgSerProValPheSerAspAsnSerSerProAlaValProGln-S 197
Db 485 CGGCACCTTCCATCGCCCGATGGGAACCGCAT-----CCTGCAGACTCTCAAGG 535
QY 197 eTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSerThrLysValProA 217
Db 536 AACACCAAGTGGTGGTGGCGCGGACACAGAGCTGTGGCAAGTCTACTCAGGTGCC 595
QY 217 laAlaTyrAlaAlaGlnGlyTyrLys---ValLeuValLeuAsnPro----- 231
Db 596 AGTACTTGTCTGGCCCGCGCTTACGTACGTGGCATGCATCAGCCCGGGAATTCGCT 655
QY 232 -----SerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHisGlyIleA 250
Db 656 GCATCTCACTGCCCAAGCGGTGGCTTCGAGAGCCTCAGTCAGTATGGTCCCGAGTTG 715
QY 250 spProAsnIleArgThrGlyValArgThrIleThrThrGlySerProIleThrTyrSerT 270
Db 716 GCTACCAAGATCCGCTTT---GAAAGCACTCGTGGCAGGACCACCAAGATCGTGTCTCTGA 772
QY 270 hrTyrGlyLysPheLeu-----AlaAspGlyGlyCysSerGlyValaLysAspIleI 288
Db 773 CGGTGGGCTCTCTCCTACGCCAGATCAGCGTGCAGCCAGCTCCCGGAGTACAGGTCC 832
QY 288 leIleCysAspGluCysHisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValL 308

```



```

Db      385 GGTGTGAAGCTTGGGAACGAGGTGGCTACAGATCCGGTTTGAGACTGCACCTCAGAG 444
QY      258 ArgThrIleThrThrGlySerProIleThrTyrSerThrTyrGly-----Lys 273
      |||||:::
Db      445 CGAACTGTT-----CTCCGTACATCAGACATGGAATGCTACCTCCGAGAG 489
QY      274 PheLeuAlaaspGlyGlyCysSerGlyGlyAlaTyrAspIleIleCysaspGluCys 293
      |||||:::
Db      490 TTCCTCTCTGAGCTGACCTTGCA-----AGTTACAGTGTGTGATGTTGATGAAGCT 543
QY      294 HisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThr 313
      |||||:::
Db      544 CACGACGGACCTTGACACAGACATCTCTTGGATTGATCAAGACGTCGCTAGATTC 603
QY      314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
      |||||:::
Db      604 CGACCTGAGCTCAAGGTCCTGTGTGGTTCAGCCACA 639

RESULT 15
LOCUS   CD240900
DEFINITION AGENCOURT_14124904 NIH_MGC_177 Mus musculus cDNA clone
IMAGE:30380772 5', mRNA sequence.
ACCESSION CD240900
VERSION   CD240900.1 GI:31001364
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 738)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM173 row: f column: 13
High quality sequence stop: 482.

FEATURES
Location/Qualifiers
1..738
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30380772"
/lab_host="DH10B (Ti-phage-resistant)"
/clone_lib="NIH_MGC_177"
/note="Organ: liver; Vector: pDNR-LIB; Site_1: Sfil
(ggcatatggccc); Site_2: Sfil (ggcgccctggcc); cDNA made
by oligo-dt priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAGCGAGGTGGCATTAGCGCGGG-3' and
5'-ATTCTAGGCGCGGCGGACATG-gt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of W.
Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores:
Pred. No.: 0.414 Length: 738
Score: 127.50 Matches: 45
Percent Similarity: 44.19% Conservative: 31
Best Local Similarity: 26.16% Mismatches: 69
Query Match: 3.52% Indels: 27
DB: 14 Caps: 7

```

```

US-09-930-591-2 (1-686) x CD240900 (1-738)
QY      174 GluSerLeuGluThrThrMetArgSer---ProValPheSerAspAsnSerSerProPro 192
      |||||:::
Db      62 GAATCTATCCAGGCTGTGCGCGCGAGCTGCGCGTTCCTCCCTTCCGAGAGGAGCTTCTG 121
QY      193 AlaValProGlnSerTyrGlnValAlaHisLeuHisAlaProThrGlySerGlyLysSer 212
      |||||:::
Db      122 GCCGCCATTGCAACACCATCAGGTCTCTCATCTCAAGGCGGAGACTGGCTCTGGGAAGACC 181
QY      213 ThrLysValPro-----AlaAlaTyrAlaAlaGlnGlyTyrLysValLeu 227
      |||||:::
Db      192 ACACAGATCCACAGTACCTCTTTGAGGAGGGTTACACAAAGAAGGCGCATGAAGATTGCT 241
QY      228 ValLeuAsnProSerValAlaAlaThrMetGlyPheGlyAlaTyrMetSerLysAlaHis 247
      |||||:::
Db      242 TGCACCCCGCGGAGAGTGGCGCTATGAGTGTGCAGCCCGAGTGGCGCGGAGATG 301
QY      248 Gly-----IleAspProAsnIleArg-----ThrGlyVal 257
      |||||:::
Db      302 GGTGTGAAGCTTGGGAACGAGGTGGCTACAGCATCCGGTTTGAGGACTGCACCTCAGAG 361
QY      258 ArgThrIleThrThrGlySerProIleThrTyrSerThrTyrGly-----Lys 273
      |||||:::
Db      362 CGAACTGTT-----CTCCGTACATCAGACATGGAATGCTACTCCGAGAG 406
QY      274 PheLeuAlaaspGlyGlyCysSerGlyGlyAlaTyrAspIleIleCysaspGluCys 293
      |||||:::
Db      407 TTCCTCTCTGAGCTGACCTTGCA-----AGTTACAGTGTGTGATGTTGATGAAGCT 460
QY      294 HisSerThrAspAlaThrSerIleLeuGlyIleGlyThrValLeuAspGlnAlaGluThr 313
      |||||:::
Db      461 CACGAGCGGACCTTGCACACAGACATCTCTTTGGATTGATCAAGACGTCGCTAGATTC 520
QY      314 AlaGlyAlaArgLeuThrValLeuAlaThrAlaThr 325
      |||||:::
Db      521 CGACCTGAGCTCAAGGTCCTGTGTGGTTCAGCCACA 556

```

Search completed: September 16, 2004, 00:40:41

Job time : 4609 secs

This Page Blank (uspio)